

Reproduced by



CENTRAL AIR DOCUMENTS OFFICE

WRIGHT-PATTERSON AIR FORCE BASE - DAYTON, OHIO

REEL-C

5228

A.T.I

116518

"NOTICE: When Government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the U.S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto."

UNCLASSIFIED

(UNPUBLISHED CARD) UNCLASSIFIED

AT: 116 518

(COPIES OBTAINABLE FROM CADO)

PRINCETON UNIVERSITY, STATISTICAL RESEARCH GROUP, N.J.

"STAIRCASE" METHODS OF SENSITIVITY TESTING

ANDERSON, T.W.; MC CARTHY, P.J.; TUKEY, J.W. 21 MARCH 46
135PP TABLES, GRAPHS

BUREAU OF ORDNANCE, WASH., D.C. (NAVORD REPORT 65-46)

EXPLOSIVES - TEST METHODS AND
INSTALLATIONS

ORDNANCE AND ARMAMENT (22)
TESTING (14)

EXPLOSIVES - SENSITIVITY
TESTING

UNCLASSIFIED

ORIGINAL
FILE COPY
116-5718

"STAIRCASE" METHOD
OF
SENSITIVITY TESTING



NAVY DEPARTMENT
BUREAU OF ORDNANCE
WASHINGTON 25, D.C.

Rear Admiral G. F. Hussey, USN
Chief of the Bureau of Ordnance

Captain K. H. Noble, USN
Director, Research and
Development Division

Commander F. I. Cone, USN
Ammunition and Explosives

Commander Stephen Brunauer, USNR
High Explosives and Amphibious
Weapons

Research Group on Theoretical Physics

NAVORD Report 65-46

"STAIRCASE" METHODS OF SENSITIVITY TESTING

T. W. Anderson, P. J. McCarthy, and J. W. Tukey

Statistical Research Group
Princeton University
Princeton, New Jersey

21 March 1966

NAVY DEPARTMENT
BUREAU OF ORDNANCE
WASHINGTON 25, D.C.

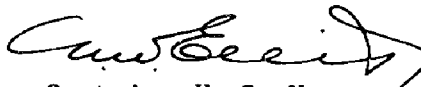
21 March 1946

NAVORD REPORT 65-46

"STAIRCASE" METHODS OF SENSITIVITY TESTING

1. NAVORD REPORT 65-46 presents information on the "Staircase" method of testing sensitivity of military explosives for the use of agencies conducting explosives research.
2. This report presents a comprehensive study. Additional information and criticism by other agencies, however, is invited.
3. The findings herein represent the interpretations of the authors and are not necessarily the views of the Bureau of Ordnance. The work has been done under contract NOrd 9240 at Princeton University by the Statistical Research Group, where it was supervised by Dr. S. S. Wilks. The BuOrd Liaison officer was Dr. R. J. Seeger.
4. This report does not supersede any existing publication.

G. F. HUSSEY, JR.
Rear Admiral, U. S. Navy


Asst. Captain, U. S. Navy
Director, Research and
Development Division
by direction

"STAIRCASE" METHODS OF SENSITIVITY TESTING

TABLE OF CONTENTS

SECTION	Page
1. Summary	1
I. GENERAL	
2. Introduction	3
3. Scope of the Present Study	5
4. Discussion of Quantitative Criteria of Efficiency	6
5. Recommended Methods	9
a. Summary	9
b. Efficiency	11
c. Simplicity	11
d. Stability	11
6. Probable Scope of Staircase Methods	31
II. TECHNICAL	
7. Introduction	32
a. Some Mathematical Preliminaries	32
b. Outline of the Investigation	35
8. The Possible Adjustments	36
9. Staircase Methods for Step Size of 0.5 Standard Deviation	45
a. General discussion	45
b. Single Explosion plus m Trials Methods	46
c. Cascade Methods	52
10. Staircase Methods at Different Step Sizes	57
a. General	57
b. Correction for step size	57

c.	The NPF and NPFI Methods	60
d.	Cascade Methods	61
e.	Single Explosion plus m Trials Methods	68
f.	Estimating the step size	73
11.	The Sequential Method	75
a.	General discussion	75
b.	Sequential Probability Ratio Plan	79
c.	Effect of changes in α and β	83
d.	Effect of point of truncation	85
e.	Effect of change in interval size	86
12.	The Pursuit Method	87
13.	The Picatinny Method	88
14.	The Up and Down Method	90
a.	Introduction	90
b.	The method of estimation	90
c.	Accuracy	91
d.	Confidence intervals	97
e.	Comparison with other methods	97
15.	Future Research in Methods	98
a.	Number of trials	98
b.	Corrections for curvature	98
c.	Confidence intervals	99
d.	Ways of testing normality	99
e.	Non-normal operation	99
f.	Block Up and Down Methods	99
g.	Variable Interval Methods	99
h.	Multiple Cascade Methods	99
i.	Other Sequential Methods	99
j.	Investigation of boundaries	100
k.	Differential effects of interval size	100
l.	Variable length-accuracy	100

SECTION	Page
m. Delayed Staircase Methods	100
n. Variable number of trials per level	100

III. COMPUTATION

16. Introduction	101
17. General Assumptions	101
18. Single Explosion Design	101
19. Cascade Methods	105
20. Single Explosion plus m Trials	111
21. NPF Inverted Method	115
22. The Picatinny Method	119
23. Sequential Method	121
24. The Pursuit Method	132

FIGURES

1. Per cent Explosions Versus Severity of Test	13
2. Average Number of Trials for the NPF Design (under the assumption that the test is started where almost no non-explosions occur)	17
3. Standard Deviation of Estimates by the Up and Down Method (with a Step Size of Approximately 1.5σ)	21
4. Standard Deviation of Estimates by the Up and Down Method (with a Step Size of Approximately $.5\sigma$)	23
5. Average Number of Trials for the Single Explosion Design (under the assumption that the test is started where almost no explosion occurs)	26
6. Average Number of Trials for the Sequential Design (under the assumption that the test is started where almost no explosion occurs)	30
7. $E(x r)$ as a Function of r for a Cascade Design ($k = 1$, $m = 1$, $h = 3$)	40
8. $E(x+ar)$ for a Single Explosion plus One Trial Design (one trial on each level, final trial one level below level of first explosion).	43
9. Average Level Estimated by the Single Explosion Design (one trial on a level)	58

FIGURES

10.	The Expected Level of the First Explosion as a Function of \bar{r} for a Cascade Design ($k = 1, m = 1, n = 1$)	62
11.	Plot Showing Relationship between Accuracy per Trial and M_{δ} for Various Cascade Designs	65
12.	Plot Showing Relationship between Accuracy per Explosion and Trials per σ' for Various Cascade Designs	67
13.	Plot Showing Relationship Between Accuracy per Trial and M_{δ} for Various Single Explosion plus m Trials Designs	70
14.	Plot Showing Relationship Between Accuracy per Explosion and Trials per σ' for Various Single Explosion plus m Trials Designs	71
15.	Expected Value of Total Deviation (about Median) in Interval Sizes for 5 Single Explosion (one trial on each level) Tests.	74
16.	Correction to be Subtracted from Average Level of First Explosion in order to Estimate the 10 Per cent Point	76
17.	Probability of Moving to Level $x + \delta$ After Making Ten Trials on Level x	78
18.	Standard Deviation of the 50 per cent Point Estimated by the Up and Down Method	92
19.	Standard Deviation of the Estimate of the Population σ' by the Up and Down Method	93
20.	Accuracy per Trial for Up and Down Method (Interval size of $1.5\sigma'$)	94

"STAIRCASE" METHODS OF SENSITIVITY TESTING

1. Summary.

In this report we are concerned with a class of methods used for determining, on the basis of a number of trials; the level of severity at which a fixed percentage of samples of a particular explosive will explode. A general discussion of the problems of sensitivity testing is given in Sections 2, 3, and 4. The methods considered can be broadly defined as follows:

A "staircase" method is any method where the severity of the next trial or group of trials is directly determined by the results of the last trial or group of trials.

Beyond the fact that all the methods considered in this report are "staircase", we may further classify them as:

- (i). Methods which have not been used before, such as the Single Explosion plus m Trials Methods, the Cascade Methods and the Sequential Method.
- (ii). Methods which have been used before, but for which an adequate analysis has not been available, such as the Naval Powder Factory and the Picatinny Methods.
- (iii). Methods which have been used before with an adequate analysis, such as the Up and Down Method.*

In general, these methods (with the exception of the Up and Down) require a varying number of trials. Consequently they can best be applied when

- (i). trials are to be made one after another,
- (ii). the result of a trial is immediately available, and
- (iii). changes in severity are easy to make.

The operation of each method on a particular sensitivity curve can be summarized by four numbers, namely, the percentage point estimated on the average, the variance of the estimated percentage point, the average number of trials required

* "Statistical Analysis for a New Procedure in Sensitivity Experiments", by Princeton University Statistical Research Group. Report No. 101.1R of the Applied Mathematics Panel, NDRC, July, 1944.

for one determination of the percentage point and the average number of explosions required for one such determination. For the methods discussed in this report, these quantities have been computed numerically, assuming that the relationship between the level of severity and the per cent explosions can be represented by a cumulative normal curve. The results of these computations are summarized in the Technical Part while the details of the computations are explained in the Computational Part.

Upon the basis of these computations, we have given in Section 5 a list of seven recommended methods, each method being accompanied by an operating procedure and an appropriate analysis of results. These recommended methods, taken singly or in groups, can be used to estimate the 10 per cent, 50 per cent or 90 per cent points with certain minimum properties. Thus we can choose a method to estimate the 10 per cent point which has one of the following four properties:

- (i). Uses a minimum number of trials,
- (ii). Uses a minimum number of explosions,
- (iii). Minimizes the assumption of normality, or
- (iv). Attempts to minimize both the number of trials and the assumption of normality (and, therefore, usually minimizes neither).

Complete recommendations are summarized in Table 1.

Section 6 gives an estimate of the degree of improvement which one might expect to obtain by future research in "staircase" methods while Section 15 points out certain obvious directions in which this research might proceed.

I. GENERAL

2. Introduction.

The term "sensitivity test" is commonly applied to tests in which an increasing percentage of individuals fail, die or explode as the severity of the test is increased. In this report we shall always speak of a sample of explosive as either exploding or failing to explode when it is subjected to a certain severity test. Partial explosions are to be classified as explosions or non-explosions according to any fixed rule. However, it should be understood that the results apply equally well to situations where individuals either die or fail to die when subjected to a certain severity test. In the first instance the severity test is usually a weight being dropped from a specified height, while in the second it may consist of the administration of a specified dose of a drug to an experimental animal. The problem of designing and analyzing such tests is a statistical one, since the determination of a percentage by repeated tests is a statistical one. If the test is properly conducted, the sample tested gives a fair representation of the explosive being studied, but this does not mean that the observed fraction of explosions is equal to the true fraction in question. The present report considers a particular group of experimental designs, determining effective and complete methods of analysis for a selected few and comparing their efficiency on several bases.

The aim of a sensitivity test is to provide estimates of one or more numbers which describe the way in which the percentage "exploding" increases with the level of severity of the test. The choice of quantities to be estimated normally involves striking a balance between what is desired and what is attainable. The quantities which are frequently desired are:

- (1). an estimate of the percentage exploding at a fixed severity, and
- (2). an estimate of the severity at which a fixed percentage will explode.

The quantities which are usually obtainable are:

- (1). an estimate of the constants which complete the specification of the increase, assuming a particular, simple form for the increase of percentage exploding with severity,

- (ii). an estimate of the severity at which a fixed percentage explode, assuming a particular, simple form for the increase of percentage exploding with severity, and
- (iii). an estimate of the percentage exploding at a fixed severity, provided this percentage is sufficiently different from 0 per cent or 100 per cent.

When an estimate of the percentage exploding at a fixed severity is required, and when this percentage is neither very small nor very large, the testing problem is very simple. It is merely necessary to make an adequate number of tests at this fixed level.

When the percentage exploding at the fixed severity is very small or very large, it is usually not feasible to make enough tests at this severity to obtain a useful estimate of the percentage, and some other device must be used. All known methods depend on the assumption of a specific functional relationship involving several constants for the increase of percentage exploding with severity, -- this dependence is important and there seems to be no way of avoiding it.

When an estimate of the severity at which a fixed percentage will explode is desired, and when this percentage is neither very small nor very large, the testing problem can be handled with relative ease. Various methods are available, and the only complications arise from the need of balancing the number of tests for a given accuracy against the extent of dependence of the estimate on the particular form assumed for the increase of percentage exploding with severity.

When an estimate of the severity at which a very small or very large percentage will explode is required, the problem is just as difficult, and the answer is just as unsatisfactory as in the case of estimating a very small or very large percentage at a fixed severity, and for the same reason.

In summary, therefore, there are three levels of dependence on assumptions:

A. No dependence on assumed form.

Estimation of moderate percentage (i. e., neither very small nor very large) at fixed severity.

B. Usually unimportant dependence on assumed form.

Estimation of severity corresponding to moderate percentage.

C. Important dependence on assumed form.

Estimation of severity corresponding to extreme percentage.

Estimation of extreme percentage at fixed severity.

3. Scope of the Present Study.

The methods discussed in this report are intended to estimate the severity at which a moderate percentage will explode assuming that severity can be measured on a scale for which the percentage exploding varies with severity according to a cumulative normal distribution (see Figure 1). We shall use the term levels to denote equally spaced positions on this scale. This normal curve could be replaced by some other curve, and may have to be as a result of future research, but there seems to be no possibility of avoiding some choice of scale. As the general discussion above predicts, the dependence of the estimate on the assumption is only moderate. However, some attention has been paid to the extent to which deviations from normality affect the various methods.

The term staircase method is applied to any method where the severity of the next trial or group of trials is directly determined by the results of the last trial or group of trials. Four such staircase methods are described briefly below as examples:

The NPF Inverted Design (Naval Powder Factory).

Starting at a level at which almost no explosions are expected, step up one level after each non-explosion. When an explosion occurs, step down one level and start to make a group of three trials. If all three fail to explode the test is concluded. When an explosion occurs, move down one level and start a new group of three, proceeding according to the same rules.

One Possible Cascade Design.

Starting at a level at which almost no explosions are expected, step up one level after each non-explosion. When an explosion occurs, step down 3 levels and start again. Step up one level after each non-explosion and stop the test after the next explosion.

Another Cascade Design.

Starting at a level at which almost no explosions are expected, make groups of two trials, stepping up one level after each pair of non-explosions. When an explosion occurs, step down one level and start again in pairs. Stop the test after the next explosion.

The Up and Down Design.

Starting at a level where about 50 per cent explosions are expected, move down one level after each explosion and up one level after each non-explosion. Stop the test after an assigned number of trials.

The present report makes a more or less complete study of a considerable number of staircase methods and presents (in Section 5) a set of recommended methods, with detailed directions and methods of analysis. The recommendations are based on the extent to which the tests attain accuracy while minimizing

- (a). the number of trials required, or
- (b). the number of explosions required, or
- (c). the sum of the number of explosions and one-tenth the number of trials.

Some attention has also been paid to simplicity of operation and analysis.

Criteria (b) and (c) above are pertinent in tests where the occurrence of an explosion is more destructive or time-consuming than the occurrence of a non-explosion. The factor 1/10 was chosen arbitrarily. In some analyses, it might be desirable to choose a different value to achieve a proper balance between explosions and non-explosions.

While this report is not an exhaustive study of staircase methods, it does outline the possibilities of such methods (as explained in Section 6).

4. Discussion of Quantitative Criteria of Efficiency.

A frequent situation in sensitivity testing is the following:

We are prepared to make N trials per test on the average. We desire as good estimates as possible.

Since the number of trials per test varies from test to test in many of the staircase methods, the present analysis measures the labor involved by the average number of trials. The number of trials usually depends very markedly on the interval size used, and thus the same method, if used once, will give different accuracies and different average numbers of trials at different interval sizes. Because of the relationship between number of trials, accuracy and interval size, the same high accuracy can often be obtained with about the same number of trials by:

- (1). using a very small interval size, or
- (ii). using a larger interval size and repeating the same method two or more times on each sample.

In (11) the estimated percentage point is taken as the average of the separate determinations.

If N trials per sample is a definite requirement for each sample (for example, when the samples require careful advance preparation), and if N is small (< 25), then the situation requires careful investigation beyond the scope of the present report. In all other cases, however, and it seems likely that this includes most of the cases of practical importance, it will usually be enough to characterize the efficiency of the test in obtaining accuracy from few trials by the "Accuracy per trial" which is calculated as follows:

$$\text{Accuracy per trial} = \frac{1}{\left(\begin{array}{c} \text{mean square} \\ \text{error of a} \\ \text{single test} \end{array} \right) \times \left(\begin{array}{c} \text{average number} \\ \text{of trials} \\ \text{per test} \end{array} \right)},$$

where the error is measured in units of the standard deviation of the (assumed) normal distribution.

When the other criteria apply, there is rarely any fixed limitation on the number of trials and the natural criteria are the

$$\text{Accuracy per explosion} = 1 / \left[\left(\begin{array}{c} \text{mean square error of a single test} \end{array} \right) \times \left(\begin{array}{c} \text{average number of explosions per test} \end{array} \right) \right]$$

and the

$$\text{Weighted Accuracy} = 1 / \left[\left(\begin{array}{c} \text{mean square error of a single test} \end{array} \right) \times \left(\begin{array}{c} \text{average number of explosions per test} + \\ (1/10) \text{ average number of trials per test} \end{array} \right) \right]$$

To explain and partly justify these criteria, consider the case of an agency which is willing to make 100 trials on a specific sample and which has to choose between

$$\text{Method A.} \quad \left\{ \begin{array}{l} \text{mean square error} = 0.3, \\ \text{average number of trials} = 10, \end{array} \right.$$

8.

$$\text{Method B.} \quad \begin{cases} \text{mean square error} = 0.5, \\ \text{average number of trials} = 5. \end{cases}$$

If method A is used, 100 trials will allow about 10 repetitions, and the mean square error of the result will be about

$$\frac{0.3}{10} = 0.03.$$

If method B is used, 100 trials will allow about 20 repetitions, and the mean square error of the result will be about

$$\frac{0.5}{20} = 0.025.$$

The approximate mean square errors are each given by $1/[100 \times (\text{accuracy per trial})]$. Clearly method B is to be preferred.

Suppose now that an agency is willing to spend a week testing a sample, and finds that it can make 15 trials a day if none are explosions, and only 5 a day if one is an explosion. Thus one explosion requires the same time as 11 non-explosions. Then if the agency is comparing

$$\text{Method A.} \quad \begin{cases} \text{mean square error} = 0.3 \\ \text{average number of trials} = 10 \\ \text{average number of explosions} = 2 \end{cases}$$

with

$$\text{Method B.} \quad \begin{cases} \text{mean square error} = 0.5 \\ \text{average number of trials} = 5 \\ \text{average number of explosions} = 1.3 \end{cases}$$

it would be natural to calculate as follows. One test using method A would require $(8 + 2 \times 11)/15 = 2$ days on the average. In 6 days, there would be about $6/2 = 3$ repetitions, and the mean square error would be about

$$\frac{0.3}{3} = 0.1.$$

One test using method B would require $(3.7 + 1.3 \times 11)/15 = 1.2$ days on the average. In 6 days, there would be about $6/1.2 = 5$ repetitions, and the mean square error

would be about

$$\frac{0.5}{5} = 0.1.$$

In this instance the two weighted accuracies are identical, both being equal to 1/.9. The agency can use either method.

Similar considerations would apply concerning accuracy per explosion if the agency were only interested in the number of explosions.

5. Recommended Methods.

a. Summary. Seven methods of sensitivity testing and appropriate methods of analysis are described here, and the conditions under which their use is desirable are indicated. The methods presented have been selected on the basis of a number of considerations, the most important of which are believed to be efficiency, simplicity, and stability. A brief discussion of these is given.

The choice of a method for actual use will depend on the end point which is to be used as a measure of sensitivity, as well as on other considerations. The methods outlined here permit the use of the 10 per cent, 50 per cent, or 90 per cent point as end points.

It should be understood that the methods presented are those which appear most desirable in the present state of our knowledge. It is possible that further investigation may result in the development of new and better methods.

The recommended methods are as follows:

- Method 1 : Naval Powder Factory (NPF).
- Method 2 : NPF Inverted
- Method 3 : Up and Down - Large Interval size.
- Method 4 : Up and Down - Small Interval size.
- Method 5 : Single Explosion
- Method 6 : Sequential for 10 per cent Point.
- Method 7 : Sequential for 90 per cent Point.

The situations in which it seems best to use these various methods are summarized in Table 1. The numbers in the table indicate the numbers of the methods. Detailed descriptions of the methods are given in the succeeding pages.

TABLE 1.

Recommended Use of Methods* in terms of
Percentage Point(s) to be Estimated

Quantity to be Minimized	10 per cent point only	50 per cent point only	90 per cent point only	10 and 90 per cent points	10 and 50 per cent points	10, 50 and 90 per cent points
Number of trials	2	4	1	3, or 1 and 2	3	3
Number of explosions	5	-	-	-	-	-
Dependence on normality	6	4	7	6 and 7	4 and 6	4, 6, and 7
Number of trials and dependence on normality	6	4	7	1 and 2	2 and 4	1, 2, and 4

Method 1 : Naval Powder Factory (NPF)

Method 2 : NPF Inverted

Method 3 : Up and Down - Large Interval Size

Method 4 : Up and Down - Small Interval Size

Method 5 : Single Explosion

Method 6 : Sequential for 10 per cent Point

Method 7 : Sequential for 90 per cent Point

* Except for Method 3, interval sizes greater than 0.5σ are not advised. There should be at least 5 levels between the 10 per cent and 90 per cent points (except for Method 3).

b. Efficiency. Most of the methods resemble the NPF method by involving a varying number of trials. This analysis measures the labor involved by the average number of trials. The number of trials usually depends very markedly on the interval size used, and thus the same method, if used once, will give different accuracies and different average numbers of trials at different interval sizes. The accuracy per trial, most naturally measured by the reciprocal of the product of the variance and the average number of trials, is nearly constant over a range of interval size for many methods. Thus the same high accuracy can often be obtained with about the same number of trials by:

- (i). using a very small interval size, or
- (ii). using a larger interval size and repeating the same method two or more times on each sample.

In (ii) the estimated percentage point is taken as the average of the separate determinations.

In some types of sensitivity testing, a trial resulting in an explosion is much more costly than a trial resulting in a non-explosion. Here the natural measure of efficiency is the accuracy per explosion, which can be measured by the reciprocal of the product of variance and average number of explosions. Both accuracy per trial and accuracy per explosion have been used in the selection of these seven methods.

c. Simplicity. It is clearly desirable that a method should be simple to use and easily taught to unskilled or semi-skilled operators. This aspect has been considered, but it is recognized that such judgments are individual matters.

d. Stability. Sensitivity tests are often used to predict safety properties. That is, tests under conditions of 1 per cent, 5 per cent, 10 per cent, 20 per cent, or 50 per cent explosions are interpreted to apply to conditions of 0.1 per cent, 0.01 per cent or 0.001 per cent explosions. Such interpretations are always delicate and depend strongly on the way in which per cent explosions is assumed to vary with severity of test or severity of handling. Present methods of interpretation are frequently based on the assumption that, when severity is measured on a suitable scale, the per cent explosions -- severity curve (the sensitivity curve) is a cumulative normal curve. In the cases of comparison between two or more explosives, the

assumption that comparisons at the 10 per cent point are similar to those at the .01 per cent point is equivalent to an assumption that the sensitivity curves are similar.

The curve in Figure 1 represents the probability of explosion as a function of height. The unit of measurement we shall use is the standard deviation (σ). The distance (in the properly chosen scale) from the 10 per cent point to the 50 per cent point or from the 50 per cent point to the 90 per cent point is 1.28σ if the curve is the cumulative normal.

The choice of a scale on which the sensitivity curve is normal is frequently a necessity for the interpretation of the sensitivity test. The normal curve could be replaced by some other curve, and may have to be as a result of future research, but there seems to be no possibility of avoiding some choice of scale. With this in view, the fact that a method of assessing a 10 per cent point assumes a normal sensitivity curve seems to be of minor importance. However, some attention has been paid to the extent to which deviations from normality affect the various methods.

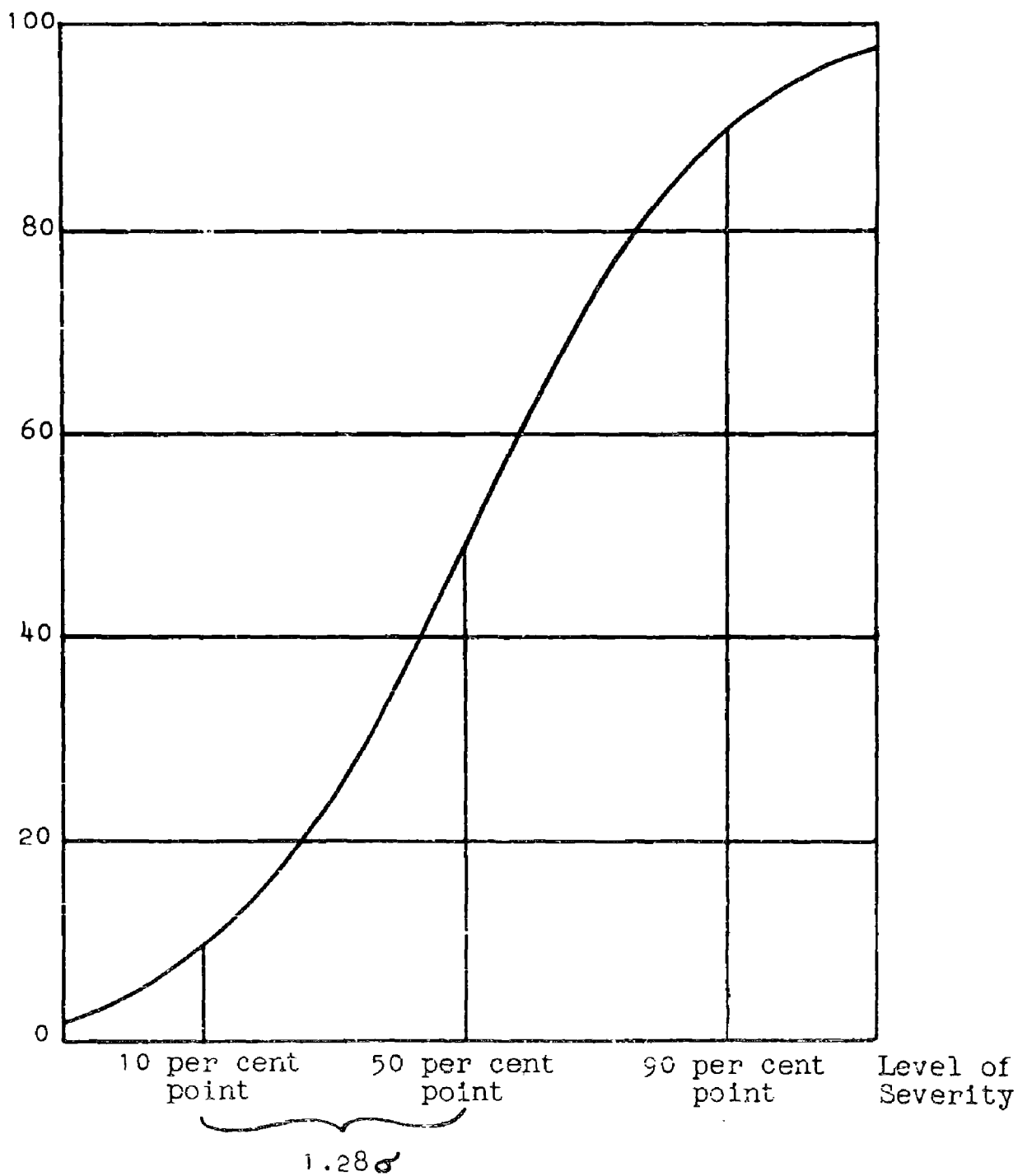
In practice, it is always advisable to plan and analyze sensitivity tests on a scale where the sensitivity curve is nearly normal. This is slightly less urgent when small interval sizes are used. The methods described below all require such a choice of scale. We shall use the term levels to denote equally spaced positions on this scale.

Since the interval size may affect the results of a sensitivity test, this aspect of stability has also been considered in selecting the recommended methods.

Figure 1

Per cent Explosions Versus Severity of Test

Per cent Explosions



Method 1 : NPF

Recommended Use: Estimation of approximate 90 per cent point when the number of trials is to be minimized.

Competing Method: For simultaneous estimation of the 10 per cent and 90 per cent points the Up and Down Method requires about the same number of trials (i. e., to obtain the same accuracy) as does the NPF (90 per cent point) plus NPF Inverted (10 per cent point), but is more dependent on assumptions.

Choice of Step and Number of Repetitions:

- (1) For maximum accuracy per trial use a step of about 0.5σ .
- (2) To control average number of trials use Figure 2.
- (3) Choose number of repetitions to obtain desired accuracy (Table 3) using average final levels.

Procedure:

- (1) Start at a level where almost all explosions are expected.
- (2) If an explosion occurs in first trial, move down one step. Repeat until a non-explosion occurs.
- (3) After the first non-explosion start moving up one step at a time as follows:

Make one trial, move up if it is a non-explosion;
 If it is an explosion make a second trial, move up if this is a non-explosion;
 If this is an explosion make a third trial, move up if it is a non-explosion;
 If an explosion on the third trial at same level occurs, end the test.
- (4) Record the level of the last test and the interval size.

Analysis A. (Rough - not recommended)

The final level estimates the 90 per cent point.

Analysis B. (Rough - adequate for less than five tests.)

- Case 1. If the step size is believed to be between 0.2σ and 0.5σ , the average final level minus .2 steps estimates the 87 per cent point.
- Case 2. If the step size is believed to be between 0.5σ and σ , the average final level minus .5 steps estimates the 84 per cent point.
- Case 3. If the step size is believed to be between 0.2σ and σ , the average final level minus .4 steps estimates the 85 per cent point. The use of this correction is less desirable than that of Case 1 or 2.

Analysis C. (Only recommended for five or more tests on the same sample of explosive involving a total of at least 75 trials.)

- (1) For a set of 5 tests add the difference between the largest and and smallest level observed to the difference between the second largest and second smallest. This sum is referred to as the Total Deviation.
- (2) Enter Table 2 with the Total Deviation expressed in interval sizes and find the correction factor (in interval sizes) for the 90 per cent point.
- (3) Add this correction to the average final level of the 5 tests.
- (4) If more than 5 tests are made (say N tests), compute the Total Deviation in a manner similar to that described above.
- (5) Multiply the Total Deviation in interval sizes by $\frac{9}{2N-1}$, enter Table 2, and proceed as above.

Accuracy: The standard deviation of the estimate is given in Table 3.

TABLE 2.

Corrections for Various Values of Total Deviation

Total Deviation in Steps for Samples of 5	Correction in Steps
3	-.1
4	+.2
5	+.4
6	+.6
7	+.6
8	+.6
9	+.6
10	+.6
12	+.5
14	+.3
16	-.1
18	-.5
20	-1.0

TABLE 3.

Standard Deviation* of Estimates by NPF Method

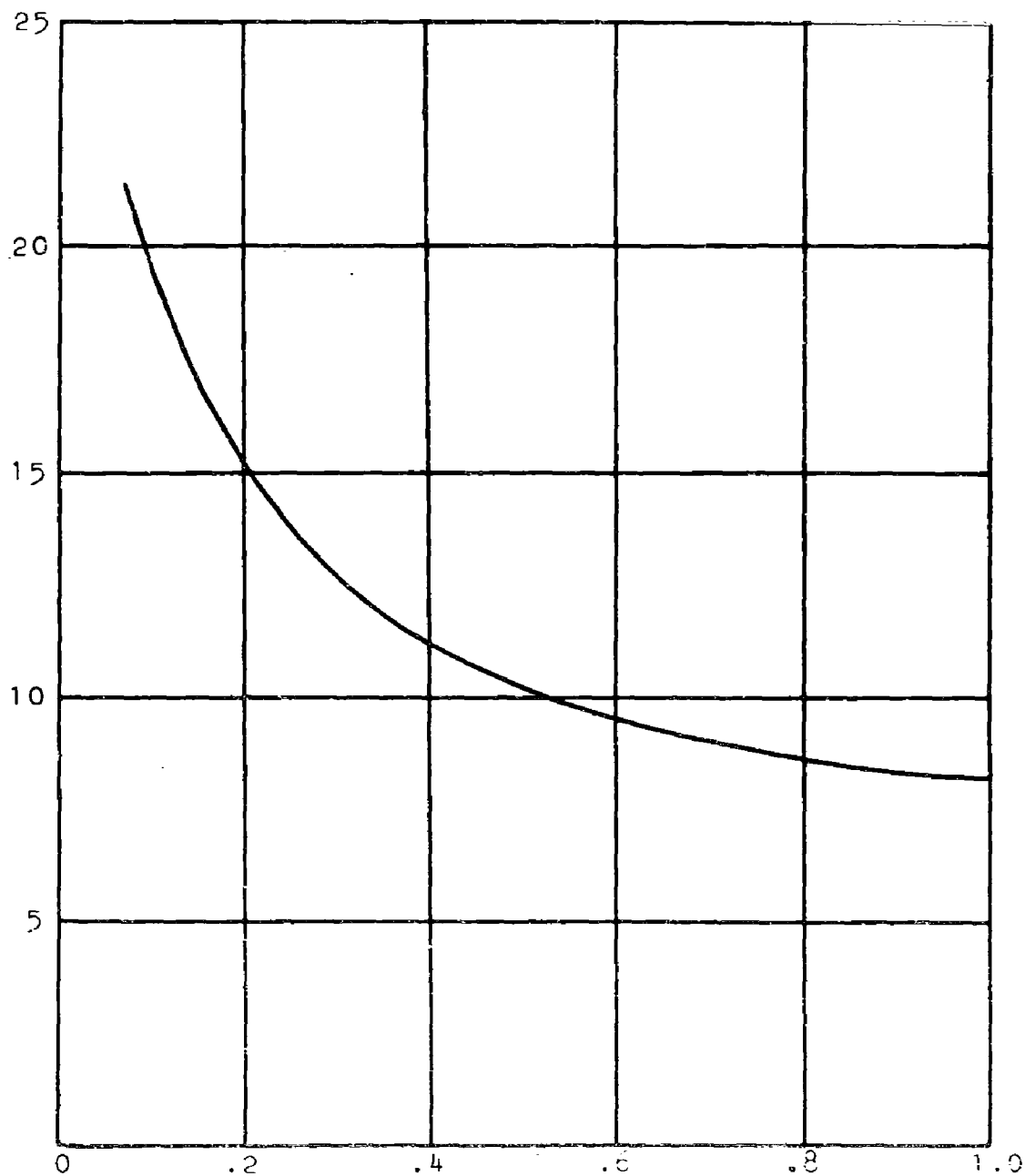
Interval Size	Single Test Standard Deviation	2 Tests Standard Deviation	5 Tests Standard Deviation	N Tests Standard Deviation
.2 σ	.50	.35	.22	$\frac{.500}{\sqrt{N}}$
.5 σ	.56	.40	.26	$\frac{.564}{\sqrt{N}}$
1.0 σ	.64	.46	.29	$\frac{.644}{\sqrt{N}}$

* The standard deviations of estimates are given in terms of σ (of the underlying distribution). 66 per cent of the estimates will fall within one standard deviation of the per cent point estimated.

Figure 2

Average Number of Trials for the NPF Design
(under the assumption that the test is started where
almost no non-explosions occur)

Average number of trials



Interval size (in σ)

Method 2 : NPF Inverted

Recommended Use: Estimation of approximate 10 per cent point when the number of trials is to be minimized.

Competing Methods: For simultaneous estimation of the 10 per cent and 90 per cent points the Up and Down Method requires about the same number of trials (i. e., to obtain the same accuracy) as does the NPF (90 per cent point) plus NPF Inverted (10 per cent point), but is more dependent on assumptions. The Single Explosion Method requires about the same number of trials, but is more dependent on assumptions.

Choice of Step and Number of Repetitions:

- (1) For maximum accuracy per trial use a step of about 0.5 σ .
- (2) For maximum accuracy per explosion use as small a step as feasible.
- (3) To control average number of trials per repetition use Figure 2.
- (4) Choose number of repetitions to obtain desired accuracy (Table 3) using average of the final levels.

Procedure:

- (1) Start at a level where almost no explosions are expected.
- (2) If a non-explosion occurs in the first trial, move up one step. Repeat until an explosion occurs.
- (3) After the first explosion start moving down one step at a time as follows:

Make one trial, move down if an explosion occurs;

If no explosion occurs make a second trial, move down if this is an explosion;

If no explosion occurs on the third trial at the same level, end the test.
- (4) Record the level of the last test and the interval size.

Analysis A. (Rough - not recommended)

The final level estimates the 10 per cent point.

Analysis B. (Rough - adequate for less than five tests.)

Case 1. If the step size is believed to be between 0.2σ and 0.5σ , the average final level plus .2 steps estimates the 13 per cent point.

Case 2. If the step size is believed to be between 0.5σ and σ , the average final level plus .5 steps estimates the 16 per cent point.

Case 3. If the step size is believed to be between 0.2σ and σ , the average final level plus .4 steps estimates the 15 per cent point. The use of this correction is less desirable than that of Case 1 or Case 2.

Analysis C. (Only recommended for five or more tests on the same sample of explosive involving a total of at least 75 trials).

- (1) For a set of 5 tests add the difference between the largest and smallest levels observed to the difference between the second largest and second smallest. This sum is referred to as the Total Deviation.
- (2) Enter Table 2 with this number (Total Deviation) expressed in interval sizes and find the correction factor (in interval sizes) for the 10 per cent point.
- (3) Subtract this correction from the average final level of the 5 tests.
- (4) If more than 5 tests are made (say N tests) compute the Total Deviation in a manner similar to that described above.
- (5) Multiply the Total Deviation in interval sizes by $\frac{9}{2N-1}$, enter Table 2, and proceed as above.

Accuracy: The standard deviation of the estimate is given in Table 3.

Method 3 : Up and Down - Large Interval Size

Recommended Use: To estimate simultaneously more than one of the 10 per cent, 50 per cent, and 90 per cent points when the number of trials is to be minimized.

Competing Methods: The NPF and the inverted NPF methods, which together estimate the 90 per cent and 10 per cent points with as small a number of trials, depend less on the assumptions.

Choice of Step and Number of Trials:

- (1) Use a step of about 1.5σ and no larger.
- (2) Choose the number of trials to obtain the desired accuracy by consulting Figure 3.

Procedure:

- (1) Start at a level near the 50 per cent point.
- (2) If the first trial results in an explosion move down one step for the next trial; if the first trial results in a non-explosion move up one step for the next trial.
- (3) After each explosion move down a step; after each non-explosion move up a step.
- (4) Record the number of explosions and non-explosions at each level.

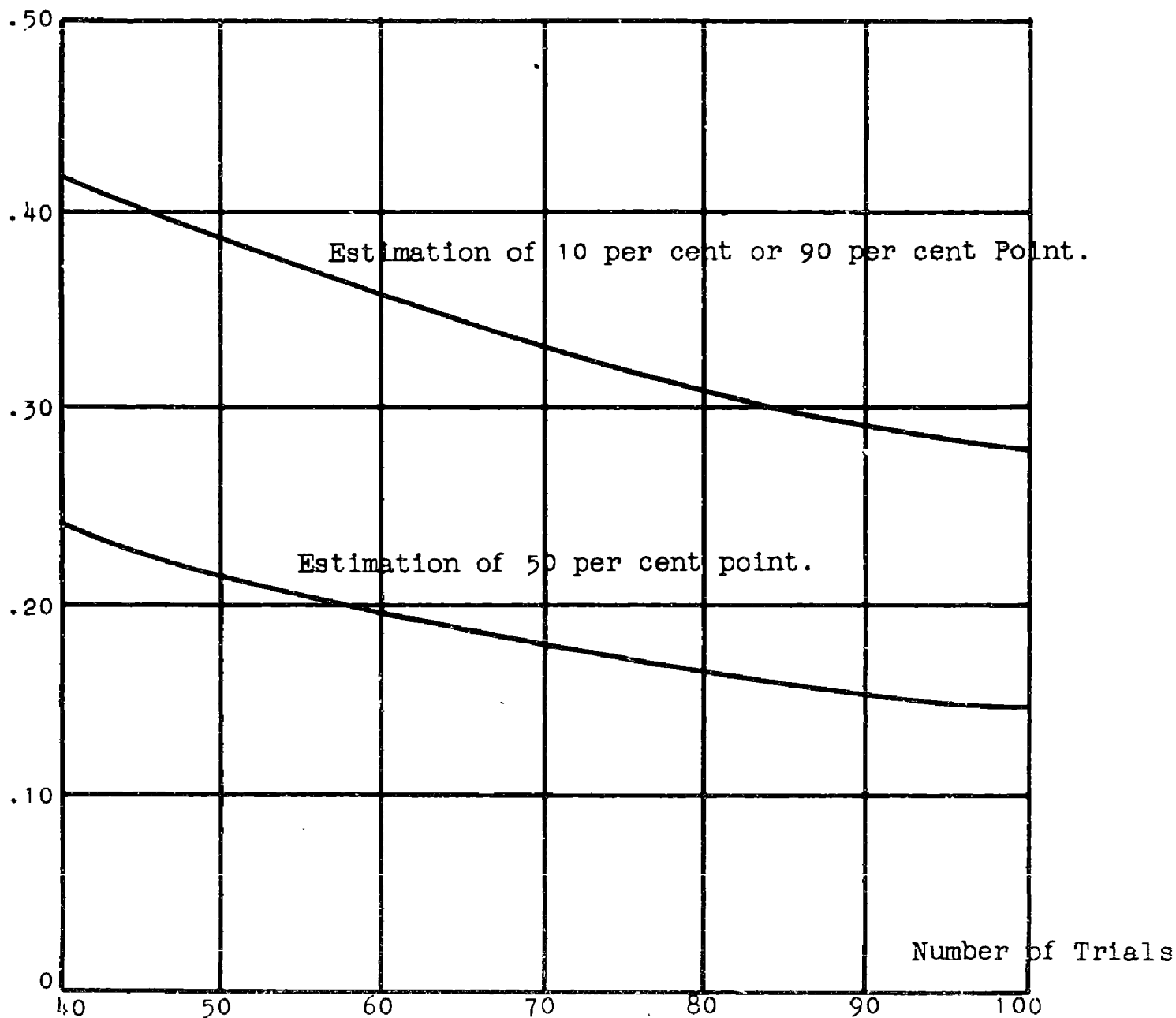
Analysis: Use the method of AMP Report No. 101.1R to estimate the 50 per cent point (m) and σ . Then $m + 1.28\sigma$ estimates the 90 per cent point and $m - 1.28\sigma$ estimates the 10 per cent point.

Accuracy: The standard deviation of the estimated per cent point is indicated in Figure 3.

Figure 3

Standard Deviation of Estimates by the
Up and Down Method with a Step Size
of Approximately 1.5σ

Standard Deviation



The standard deviations of estimates are given in terms of σ (of the underlying distribution). 66 per cent of the estimates will fall within one standard deviation of the per cent point estimated.

Method 4 : Up and Down - Small Interval Size

Recommended Use: To estimate the 50 per cent point.

Choice of Step and Number of trials:

- (1) Use a step size of about .5 or smaller.
- (2) Choose the number of trials to obtain the desired accuracy by consulting Figure 4.

Procedure:

- (1) Start at a level near the 50 per cent point.
- (2) If the first trial results in an explosion move down one step for the next trial; if the first trial results in a non-explosion move up one step for the next trial.
- (3) After each explosion move down a step; after each non-explosion move up a step.
- (4) Record the number of explosions and non-explosions at each level.

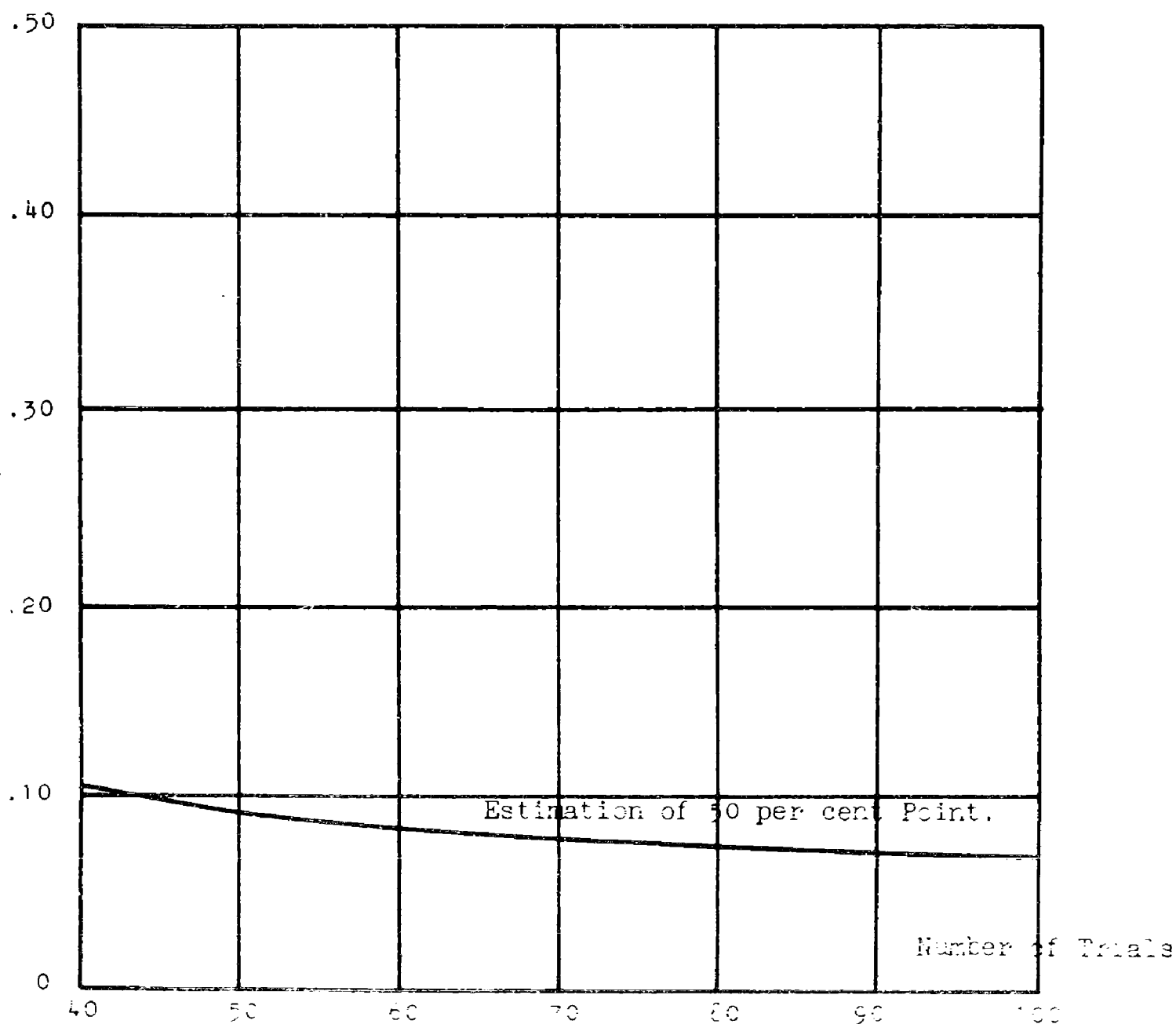
Analysis: Use the method of AMP Report No. 101.1R to estimate the 50 per cent point.

Accuracy: The standard deviation of the 50 per cent point is indicated in Figure 4.

Figure 4

Standard Deviation of Estimates by the
Up and Down Method with a Step Size
of Approximately $.5\sigma$

Standard Deviation



The standard deviation of estimates are given in terms of σ (of the underlying distribution). 66 per cent of the estimates will fall within one standard deviation of the per cent point estimated.

Method 5 : Single Explosion

Recommended Use: Estimation of low percentage point when number of explosions is to be minimized.

Choice of Step and Number of Repetitions:

- (1) For maximum accuracy per explosion use the smallest step that is feasible.
- (2) For maximum accuracy per trial use a step of about .5 σ .
- (3) To control the average number of trials per repetition use Figure 5.
- (4) Choose the number of repetitions to obtain the desired accuracy (Table 5) using the average of the final levels.

Procedure:

- (1) Start at a level where almost no explosions are expected.
- (2) If no explosion occurs on the first trial move up one step and make another trial.
- (3) Continue to move up after each non-explosion until an explosion occurs.
- (4) Record the level at which the explosion occurs and the interval size.

Analysis A. (Rough)

Case 1. If the step size is believed to be between .1 σ and .2 σ , the average final level minus 3.5 steps estimates the 7 per cent point.

Case 2. If the step size is believed to be between .2 σ and .5 σ , the average final level minus 2 steps estimates the 10 per cent point.

Note: If the step size is believed to be between .1 σ and 1.0 σ , these instructions still hold. However, this is now a rougher approximation which should not be used unless absolutely necessary.

Analysis B. (Only recommended for five or more tests on the same sample of explosive involving a total of at least 75 trials).

- (1) For a set of 5 tests add the difference between the largest and smallest level observed to the difference between the second largest and second smallest. This sum is referred to as the Total Deviation.
- (2) Enter Table 4 with this Total Deviation expressed in interval sizes and find the correction factor (in steps) for the 10 per cent point.
- (3) Subtract this correction from the average final level of the 5 tests.
- (4) If more than 5 tests are made (say N tests), compute the Total Deviation in a manner similar to that described above.
- (5) Multiply Total Deviation in interval sizes by $\frac{9}{2N-1}$, enter Table 4, and proceed as above.

Accuracy: The standard deviation of estimate is given approximately by Table 5.

Figure 5

Average Number of Trials for the Single Explosion Design
(under the assumption that the test is started
where almost no explosions occur)

Average Number of Trials

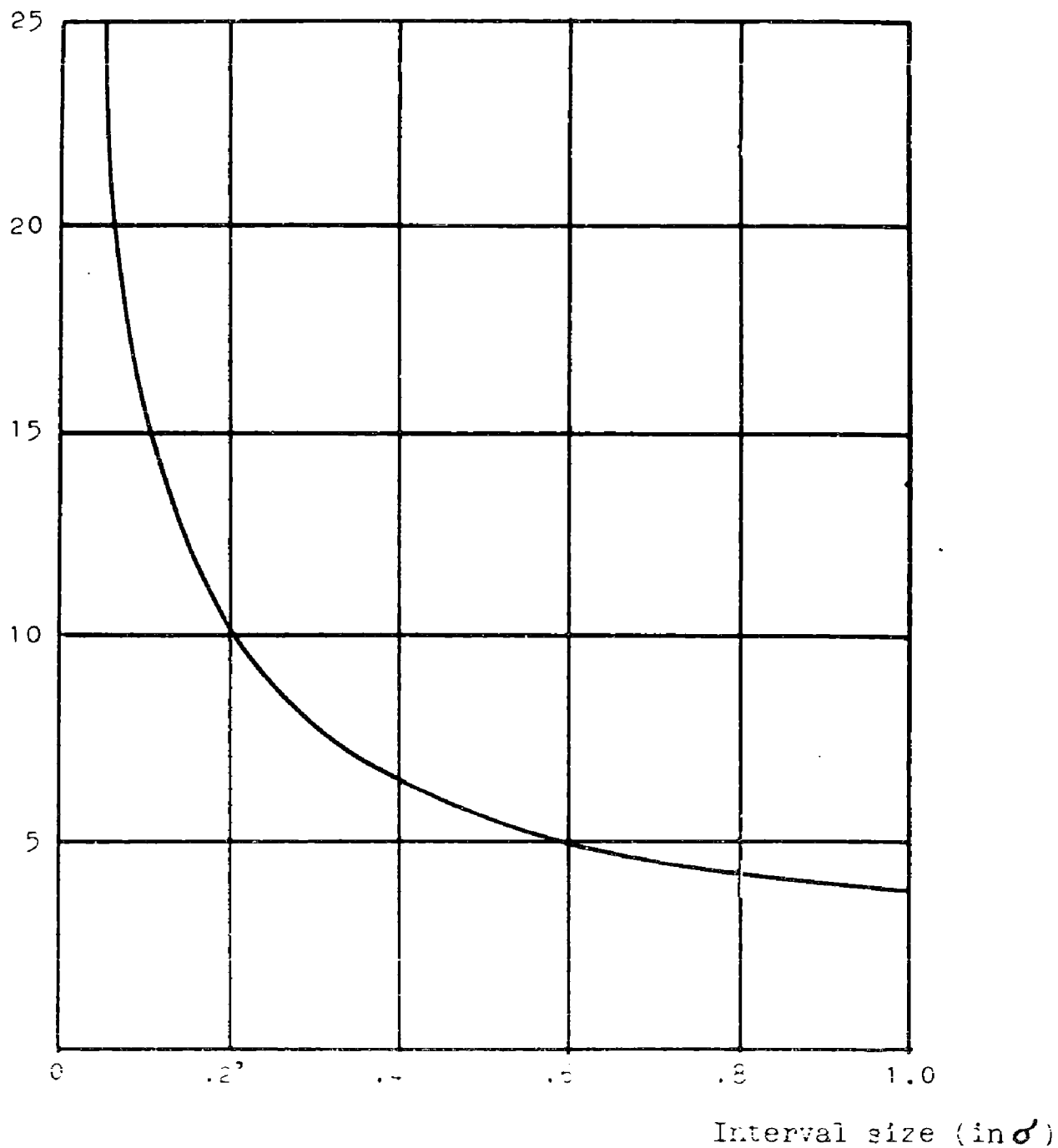


TABLE 4.

Corrections for Various Values of Total Deviation

Total Deviation in Steps for Samples of 5	Correction in Steps
3	+1.6
4	+1.8
5	+2.1
6	+2.2
7	+2.3
8	+2.4
9	+2.4
10	+2.5
12	+2.4
14	+2.1
16	+1.8
18	+1.4
20	+0.9
22	+.4
24	-.4
26	-1.3
28	-2.0
30	-2.7

TABLE 5.

Standard Deviation* of Estimates of Single Explosion Method

Interval Size	Single Test Standard Deviation	2 Tests Standard Deviation	5 Tests Standard Deviation	N Tests Standard Deviation
.2 σ	.62	.44	.36	$\frac{.622}{\sqrt{N}}$
.5 σ	.75	.53	.43	$\frac{.752}{\sqrt{N}}$
1.0 σ	.89	.63	.51	$\frac{.891}{\sqrt{N}}$

- * The standard deviations of estimates are given in terms of σ (of the underlying distribution). 66 per cent of the estimates will fall within one standard deviation of the per cent point estimated.

Method 6 : Sequential for 12 per cent Point

Method 7 : Sequential for 88 per cent Point

Recommended Use: Estimation of approximate 12 per cent point when one desires to make as few assumptions as possible concerning the underlying distribution. The recorded level estimates the 12 per cent point regardless of step size. This method may also be used to estimate the 88 per cent point if throughout the procedure non-explosion is substituted for explosion, up for down and down for up.

Competing Method: The NPF Inverted requires a smaller number of trials for the same accuracy but depends more upon the assumption concerning the underlying distribution. The two methods require about the same number of explosions.

Choice of Step:

- (1) For maximum accuracy per trial use a step of about .5σ.
- (2) To reduce the number of explosions use as small a step as feasible.
- (3) To control the average number of trials use Figure 6.

Procedure:

- (1) Start at a level where almost no explosions are expected.
- (2) If no explosion is obtained in two trials, move up one step. Continue to move up after each pair of trials until the first explosion occurs.
- (3) After the first explosion, continue to test at this level using the following procedure (disregarding the tests already made):
 - a. If two explosions are obtained out of 2, 3, 4, or 5 trials move down one step as soon as the second explosion is obtained.
 - b. If three explosions are obtained out of 7, 8, 9, 10, 11, 12 or 13 trials move down one step as soon as the third explosion occurs.

- c. If thirteen trials are made without obtaining an explosion move up one step.
 - d. If no move (as indicated by a, b or c) has been made at the end of thirteen trials, discontinue testing.
- (4) As long as no decision of type 3(d) is obtained, continue to move up or down as indicated by 3(a), 3(b) or 3(c). Discontinue testing when a decision of type 3(c) follows a decision of type 3(a) or 3(b), or when a decision of type 3(a) or 3(b) follows a decision of type 3(c).
- (5) Record:
- a. The level at which a decision of type 3(d) has been obtained, or
 - b. The midpoint of the last two levels at which testing occurred when testing has been discontinued as in (4).

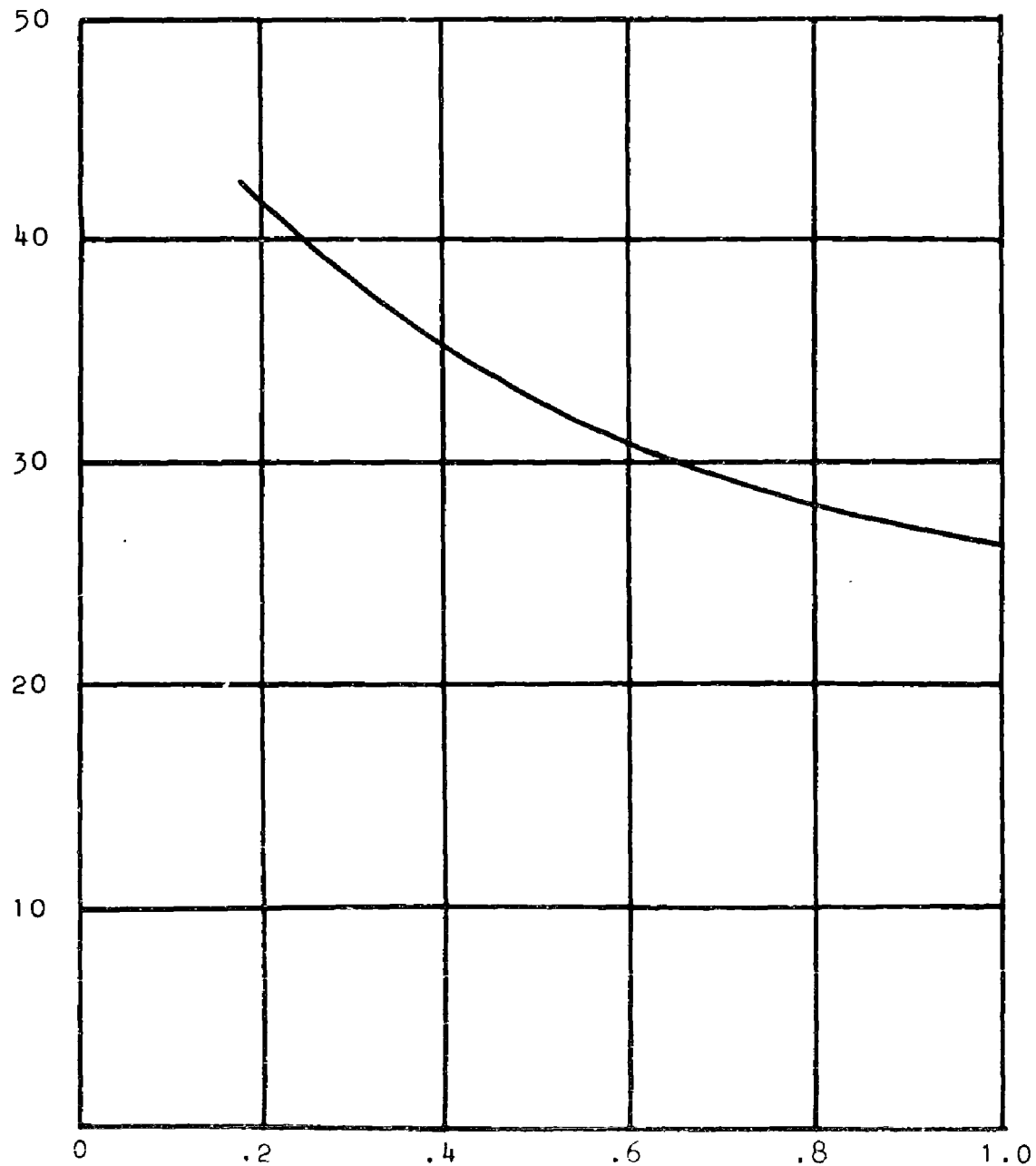
Analysis: The recorded level estimates the 12 per cent point regardless of step size.

Accuracy: The standard deviation of the estimate is approximately .40 for all step sizes.

Figure 6

Average Number of Trials for the Sequential Design
(under the assumption that the test is started
where almost no explosions occur)

Average Number of Trials



Interval size (ln σ)

6. Probable Scope of Staircase Methods.

Staircase methods are useful in testing under the conditions that

- (i). trials are to be made one after another,
- (ii). the outcome of each trial is available immediately, and
- (iii). changes in severity are easy to make.

It seems reasonable to believe that the efficiencies of the methods recommended in this report are nearly as great as those obtainable by any practical test. Curves sketched on Figures 11 to 14 show an estimate of the true boundary in comparison with the attained results. Selected points are tabulated below. The usefulness and efficiency of staircase methods in other types of sensitivity problems are uncertain, but deserve careful study.

Estimated Optimum Performance of Staircase Methods

Point Estimated	Accuracy per trial	Trials per σ^*		Accuracy per explosion
0.0 σ (50 per cent)	.60			
$\pm 1.0\sigma$ (16 and 84 per cent)	.36	2		1.90
$\pm 1.2\sigma$ (12 and 88 per cent)	.33	4		2.50
$\pm 1.4\sigma$ (8 and 92 per cent)	.29	6		2.90
$\pm 1.6\sigma$ (5 and 95 per cent)	.25	8		3.30

* The density of testing seems to determine the accuracy per explosion.

II. TECHNICAL

7. Introduction.

a. Some Mathematical Preliminaries. Expressed mathematically, the problem of sensitivity testing, in somewhat more generality than actually used here, takes the following form:

- (i). The probability of an explosion at the level x is an unknown function, p_x .
- (ii). Tests may be made at the levels $h_0 + kh$, $k = 0, \pm 1, \pm 2, \dots$ where h_0 and h have been chosen in advance on the basis of crude information or guesses about p_x . The rules for selecting successive values of k may, but need not, depend on the results of early trials.
- (iii). It is assumed that the function p_x is of the form

$$p_x = q\left(\frac{x-m}{\sigma}\right),$$

where q is a specified function and m and σ are constants depending on the explosive under test.

- (iv). From the results of the test it is desired to estimate the level x at which $p_x = \alpha$. This estimate can depend on the results of the various trials in any way.

Expressed in these terms, the basic problem is to make a good estimate with as little "trouble" as possible. Measures of goodness of estimate and amounts of trouble are discussed briefly in Part I (Section 4), with the result that we shall use the criteria given there, namely accuracy per trial, accuracy per explosion and weighted accuracy.

Of the assumptions made above, (iii) is more restrictive than it should be from the point of view of application, but, as discussed in Part I (Section 2), there seems no way to avoid it.

In actual practice we shall assume that $q(t)$ is the cumulative normal distribution with mean zero and unit variance

$$q(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{s^2}{2}} ds,$$

which is well tabled* in various forms. A graph of this function is given in Figure 1 (page 13).

We are concerned with a statistical situation and the result, for example, of making n_1 trials at $k = 3$ and n_2 trials at $k = 5$, is not certain -- it will vary from repetition to repetition.

There will often be no loss of convenience in allowing the number of trials in the resulting test to vary from repetition to repetition -- and this will often be the case in staircase tests, where the level for the next trial depends on the results of earlier trials.

Since we have a statistical situation, it will be useful to recall some elementary results. If the probability of an explosion on each of n independent trials is p , then the probability of exactly k explosions in the n trials is

$$\binom{n}{k} p^k (1-p)^{n-k},$$

where $\binom{n}{k}$ is the binomial coefficient $\frac{n!}{k!(n-k)!}$. Similarly, the probability that we obtain at least one explosion in n trials is

$$1 - (1-p)^n.$$

Another situation in which we shall be frequently interested is where we make trials (the probability of an explosion in each trial being p) until we obtain either one explosion or n non-explosions. The average number of explosions in repeated

* Fisher and Yates, Statistical Tables, Tables I and IX.

Kelley, Truman L., The Kelley Statistical Tables, Table I.

Mathematical Tables from Handbook of Chemistry and Physics, pp 200-204.

NAVORD Report No. 205-45, Tables to Facilitate the Analysis of Sensitivity Data.

Pearson, Karl, Tables for Statisticians and Biometricians, Tables I, II and III.

Work Projects Administration, Tables of Probability Functions, Vol. II.

tests of this kind is

$$1 - (1-p)^n$$

while the average number of trials is

$$1 \cdot p + 2 \cdot (1-p) p + 3(1-p)^2 p + \dots + n(1-p)^{n-1} p + n(1-p)^n = \frac{1-(1-p)^n}{p}$$

For those tests which end with an explosion, the average number of trials is

$$\left[1 \cdot p + 2 \cdot (1-p) p + 3(1-p)^2 p + \dots + n(1-p)^{n-1} p \right] / \left[1 - (1-p)^n \right] = \frac{1-(1+np)(1-p)^n}{p(1-(1-p)^n)}.$$

In many instances our test will provide two pieces of information, say x and r (see succeeding sections for the actual specification of x and r). Then the probability that we obtain particular values of x and r will be denoted by $P(x,r)$. The probability that we obtain a specified value of one of these variables, regardless of the value of the other, is given by $P(x)$ and $P(r)$ respectively, where

$$P(x) = \sum_r P(x,r), \text{ and}$$

$$P(r) = \sum_x P(x,r).$$

If we desire the probability that x has a given value when r can assume only one value, we write this as $P(x|r)$ and immediately see that this is equal to

$$\frac{P(x,r)}{P(r)}.$$

By the use of these probabilities we can define the expected (i. e., average) value of x , $E(x)$, the expected value of r , $E(r)$, and the expected value of x , given that r has a particular value, $E(x|r)$, as

$$E(x) = \sum_x x P(x),$$

$$E(r) = \sum_r r P(r), \text{ and}$$

$$E(x|r) = \sum_x x P(x|r) .$$

Now if our final reported level, L , is a function of x and r , we have that the expected value of L , $E(L)$, is

$$\begin{aligned} E(L) &= \sum_x \sum_r L(x,r) P(x,r) \\ &= \sum_x \sum_r L(x,r) P(x|r) P(r). \end{aligned}$$

Similary, the variance of $L(x,r)$ is

$$\sigma_L'^2 = \sum_x \sum_r [L(x,r) - E(L)]^2 P(x|r) P(r).$$

b. Outline of the Investigation. Various combinations of the above results by means of elementary probability theory allow us, at least in theory, to write down mathematical expressions for the criteria -- accuracy per trial, accuracy per explosion, and weighted accuracy for the methods under consideration. However, a little experience soon shows that it would be very difficult to compare the staircase methods by means of these analytic expressions. For this reason we have adopted a numerical approach to the problem. As expressed before, $q(t)$ is taken as the cumulative normal distribution for these computations with the understanding that any other $q(t)$, for which tabular values are available, could be used with equal facility. Accordingly, we have proceeded in an exploratory way, computing the three criteria for a wide variety of cases and then noting any trends which show up in the data.

Since we have no explicit a priori distribution of σ' , one of the constants appearing in the cumulative normal distribution, and no practical grounds for

assuming one, each individual computation will be conducted at a constant step size, i. e., h/σ . The usual procedure in studying a method or group of methods will be to make a complete investigation for a step size of one half the standard deviation (σ) of the underlying distribution. Having done this, we pick out the best of these methods with respect to one criterion, and then see what effect a change of step size has on these selected methods.

It should be noted that accuracy per trial, accuracy per explosion, and weighted accuracy, as defined in Section 4, are independent of the percentage point estimated, i. e., x_α . As a consequence, the "best" tests will be picked out without much attention being devoted to the average level which they estimate. After determining the "best" tests in this manner, we shall specify this average estimated level and attempt to devise adjustments which will minimize its dependence upon the step size used in making the tests.

8. The Possible Adjustments.

In addition to choosing procedures of testing, this study develops methods of analysis of the results. When a particular test is made by a staircase design, there is available at the end of the test

- (i). the result of each group of trials, and
- (ii). the results of the individual trials within the group.

Some investigation of special cases indicates that the information under (ii) is of relatively little use in increasing the accuracy of the estimate. We give it no further consideration in this report.

Let us now consider a special case of adjustment. For many of the designs considered, the information of type (i) consists of

- (a). a preliminary critical level, \underline{x} , and
- (b). a final critical level, \underline{y} .

Unless otherwise stated, \underline{x} and \underline{y} will be expressed as multiples of the standard deviation of the underlying normal distribution, the true 50 per cent point being taken as 0.00. We can do this in a theoretical study of the test, although we could not do it in an actual application. Examples are the NPF Inverted Method (the level of first explosion and the final level) and the Cascade Methods (the first and second

levels of explosion). Now suppose that we wish to estimate a percentage point from these two levels and that we are able to assume

- (i). the sensitivity curve is a cumulative normal curve, and
- (ii). some knowledge of its location and spread is at hand before the experiment is made,

where the knowledge in (ii) has probably been gained from other tests on similar samples, or from a preliminary test. It is clearly necessary to have some information with respect to (ii), and to make certain adjustments in accordance with it. Otherwise one would never know just what percentage point was being estimated. To illustrate this, let us consider one of the simplest staircase methods, namely the procedure which consists of making the first trial at a point which is, hopefully, far below the 50 per cent point and making each trial at successively higher levels until an explosion occurs. If the interval size is large (and we must have knowledge of the spread of the sensitivity curve in order to judge this), say 3 standard deviations, then the reported level will be at least the 50 per cent point (neglecting the small probability of explosion at points more than 3 standard deviations below the 50 per cent point). On the other hand, if the interval size is extremely small, say .1 standard deviation, the reported level will be much smaller.

The adjustments to be applied should accomplish two things, namely,

- (i). make the variance (for fixed δ , where δ is the interval size expressed as a multiple of the standard deviation) of the reported level a minimum, and
- (ii). make the average reported level (for fixed δ) relatively constant as δ changes.

We can make separate adjustments for these two objectives since the addition of a quantity depending only on δ will make an arbitrary adjustment to the average for fixed δ without affecting the variance for this fixed δ . In other words, we will have simply applied a translation. Let us start then by making an adjustment designed to reduce the variance for fixed δ .

For our discussion of the minimum variance we shall begin with tests which end with two critical levels, \underline{x} and \underline{y} . We immediately see that these results can also be specified by \underline{x} and \underline{r} , \underline{r} being defined as equal to $\underline{y}-\underline{x}$. \underline{r}/δ is equal to the

length of the second run (i. e., the number of levels on which trials are made) plus or minus a constant depending on the method. It is clear that the reported level should be such a function of \underline{x} and \underline{y} that the reported level for $x + a$ and $y + a$ is a more than the level for \underline{x} and \underline{y} . That is, the level is

$$L = x + f(y-x), \text{ or}$$

$$L = x + g(r).$$

If we define $P(x|r)$ as the probability of obtaining a particular value \underline{x} , given that $y-x$ has a specified value \underline{r} , and $P(r)$ as the probability of obtaining this specified value, then the average reported level, $E(L)$, is equal to

$$\sum_r \sum_x [x + g(r)] P(x|r) P(r) .$$

Similarly, the variance of this reported level, \underline{L} is

$$\begin{aligned} \sigma_L^2 &= \sum_r \sum_x [L - E(L)]^2 P(x|r) P(r) \\ &= \sum_r \sum_x [x + g(r) - E(x) - E(g(r))]^2 P(x|r) P(r) \\ &= \sum_r P(r) \sum_x [x + g(r) - E(x) - E(g(r))]^2 P(x|r). \end{aligned}$$

In adjusting the \underline{x} values, we are only permitted to add a constant amount, this constant amount being possibly different for each value of \underline{r} . Now in choosing this constant, namely $g(r)$, we are at present requiring that it should minimize σ_L^2 . It is well known* that under such circumstances, we should choose

$$\begin{aligned} g(r) &= - E(x|r) \\ &= - \sum_x x P(x|r), \end{aligned}$$

* Differentiate σ_L^2 with respect to $g(r)$, set this equal to zero, and solve for $g(r)$.

for with this choice of $g(r)$, the means of all subsets of transformed x values are equal, and consequently equal to the pooled mean. With this choice of $g(r)$ we have

$$\sigma_L^2 = \sum_r \sum_x [x - E(x|r)]^2 P(x|r) P(r)$$

since $E[E(x|r)] = E(x)$.

In order to accomplish this adjustment for r we find that if we plot $E(x|r)$ against r in terms of the standard deviation of the underlying distribution we obtain a graph that is nearly a straight line (see Figure 7). Hence we can approximate $g(r)$ by a linear function. The addition to the variance caused by this approximation is negligible.

Thus for fixed interval size, δ , we can find a constant, a , so that

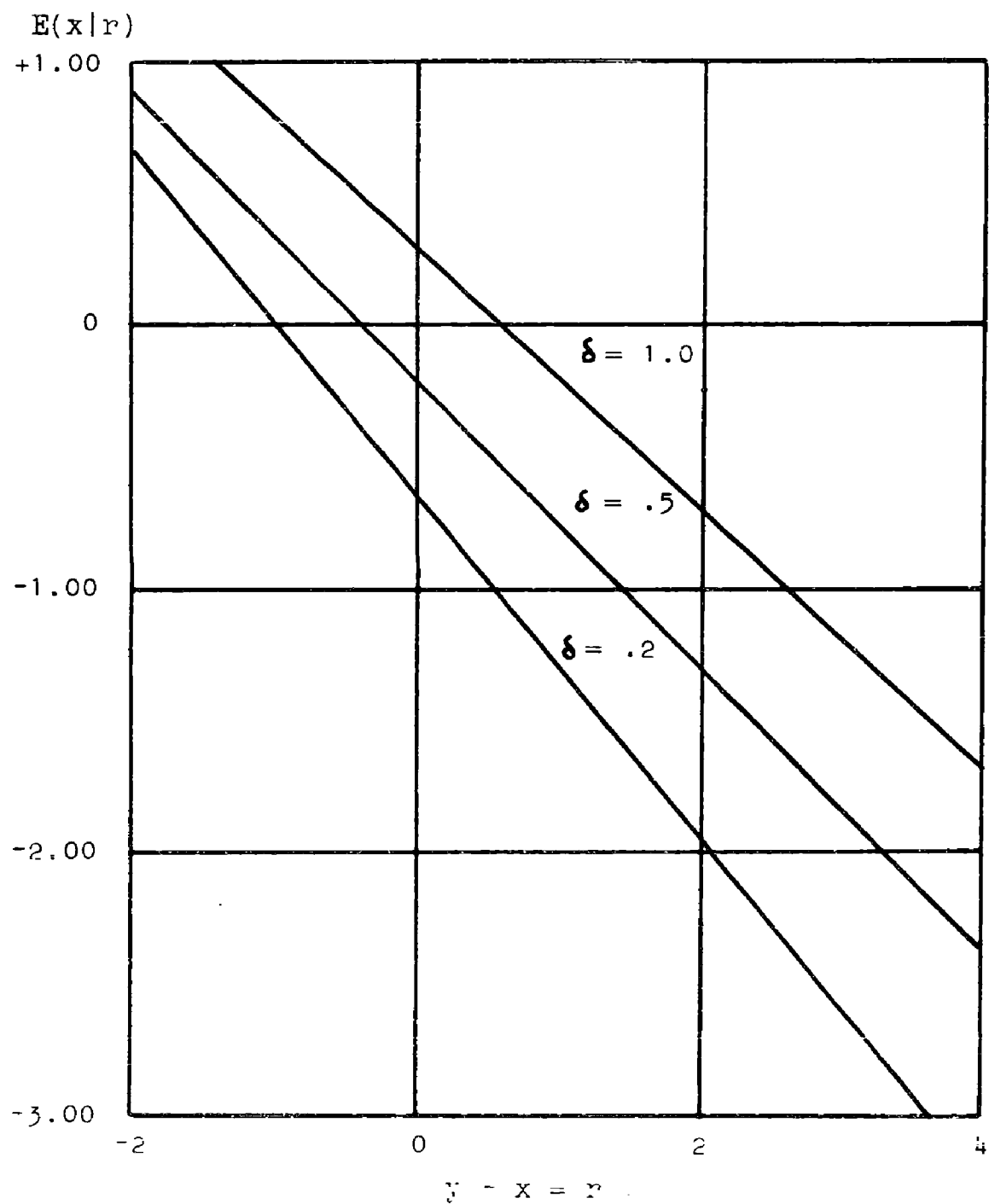
$$x + ar$$

is well adjusted for the effects of r , though not at all adjusted for the effects of possible changes in δ . The dimensionless quantity a is a function of δ , but computation shows that it varies only slowly. We shall, therefore, select an average value for each prejudged range of values of δ which we wish to consider.

To illustrate this adjustment with a very simple example of a different sort, consider the Single Explosion plus One Trial Design where one makes a single trial on each level, moving up a level after each non-explosion. This procedure is started at a level on which the probability of obtaining an explosion is almost zero. As soon as an explosion is obtained, we record this level and make one final trial on the next lower level, recording whether this trial results in an explosion or non-explosion. x is the level at which the first explosion is obtained and r is either 0 or 1 depending upon whether the final trial is a non-explosion or an explosion. Given below for several interval sizes are the expected values for x when the final trial is an explosion and when the final trial is a non-explosion.

Figure 7

$E(x|r)$ as a Function of r
 for a Cascade Design
 ($k = 1, m = 1, h = 3$)



Single Explosion plus One Trial Design

One trial on each level, final trial at one level
below the level of first explosion

Interval Size	Expected Level of first explosion when final trial is an explosion	Expected level of first explosion when final trial is a non-explosion	Variance of level of first explosion
.2σ	-.36	-.90	.39
.5σ	+.29	-.45	.56
1.0σ	+.95	-.01	.79

If the interval size is .2, then we take \underline{a} as $-.54$. This means that $E(x+ar|r=0) = E(x+ar|r=1) = -.90$. Similarly, for interval size .5, \underline{a} is $-.74$ and then $E(x+ar|r=0) = E(x+ar|r=1) = -.45$; for interval size 1.0, \underline{a} is $-.96$ and $E(x+ar|r=0) = E(x+ar|r=1) = -.01$.

Notice that in this instance the definition of r was at our disposal. If we had defined it to be 0 if the final trial were an explosion and 1 if the final trial were a non-explosion, then for interval size .2, \underline{a} would be $+.54$ and $E(x+ar|r=0) = E(x+ar|r=1) = -.36$; for interval size .5, \underline{a} would be $+.74$ and $E(x+ar|r=0) = E(x+ar|r=1) = -.29$; and for interval size 1.0, \underline{a} would be $+.96$ and $E(x+ar|r=0) = E(x+ar|r=1) = .95$.

We have thus replaced the situation tabled above by the following:

Single Explosion plus One Trial Design

One trial on each level, final trial at one level
below level of first explosion

Interval Size	\underline{a}	Expected value of $x+ar$ when final trial is an explosion	Expected value of $x+ar$ when final trial is a non-explosion	Variance of $x+ar$
0.2σ	-.54	-.90	-.90	.34
.5σ	-.74	-.45	-.45	.46
1.0σ	-.96	-.01	-.01	.60

The additional adjustment necessary for making the average reported level relatively free from choice of interval size will be based upon the fact that the addition of a multiple of the interval size will make a larger adjustment the larger the interval size. If the average reported level were a linear function of interval size, then our adjustment would make the average final reported level completely independent of interval size. Actually, the average level is not a linear function, and hence the adjustment will not be perfect.

This adjustment for changes in δ must be applied in terms of interval size and not of standard deviation since only the interval size is known accurately. Let

$$\begin{aligned}\bar{x}(\delta) &= \text{average partly corrected estimate at interval size } \delta \\ &= \text{average value of } (x + ar),\end{aligned}$$

and let b be the number of intervals to be subtracted as a correction for interval size effects. Then

$$x + ar - b\delta$$

is the final estimated level. The value of b will be chosen to make the average estimated level at interval size δ ,

$$\bar{x}(\delta) - b\delta$$

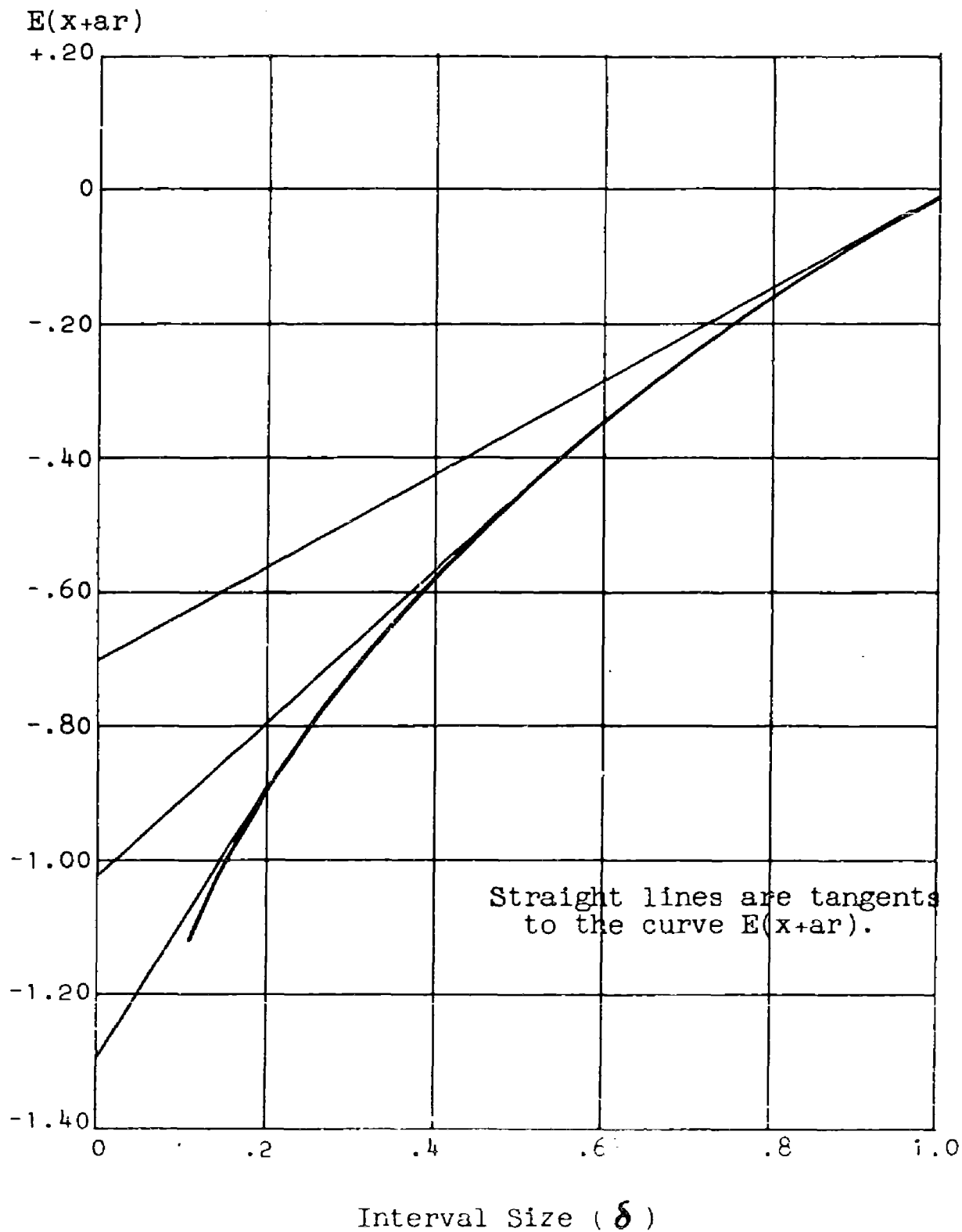
nearly constant. The best correction very near $\delta = \delta_0$ could be obtained from

$$b = \left(\frac{d\bar{x}(\delta)}{d\delta} \right)_{\delta=\delta_0}$$

This correction can best be illustrated in a graphical fashion. Thus in Figure 8 we have first plotted the average value of $x+ar$ (for interval sizes of $.2\sigma'$, $.5\sigma'$ and $1.0\sigma'$) for the Single Explosion Design which we have been considering in this section, namely one trial on each level with the final trial one level below the level of the first explosion.

Figure 8

$E(x+ar)$ for a Single Explosion plus One Trial Design
(one trial on each level, final trial one level
below level of first explosion)



Between these computed points we have then drawn a smooth curve. The value of \underline{b} for a particular δ_0 can be obtained with graphical accuracy by drawing the tangent to the curve at $\delta = \delta_0$ and taking the slope of this tangent. For this example, the tangents have been drawn at $\delta = .2\sigma$, $\delta = .5\sigma$, and $\delta = 1.0\sigma$, and the corresponding values of b are 2.1, 1.15 and .7. The results obtained by the use of these correction factors for interval size are summarized in the following table.

Application of Correction Factor for Interval Size
to a Single Explosion plus One Trial Design

(one trial on each level, final trial one level
below level of first explosion)

Interval Size (δ)	$E(x+ar)$	$E(x+ar)$ $-2.10^*\delta$	$E(x+ar)$ $-1.15^{**}\delta$	$E(x+ar)$ $-.70^{***}\delta$
.2 σ	-.90	-1.32	-1.13	-1.04
.3 σ	-.71	-1.34	-1.06	-.92
.4 σ	-.57	-1.41	-1.03	-.78
.5 σ	-.45	-1.50	-1.02	-.80
.6 σ	-.34	-1.60	-1.03	-.76
.7 σ	-.25	-1.72	-1.06	-.74
.8 σ	-.16	-1.84	-1.08	-.72
.9 σ	-.08	-1.97	-1.28	-.71
1.0 σ	-.01	-2.11	-1.32	-.71
1.1 σ	+.06	-2.25	-1.20	-.71
1.2 σ	+.12	-2.41	-1.26	-.72

* Slope of tangent drawn at .2 σ

** Slope of tangent drawn at .5 σ

*** Slope of tangent drawn at 1.0 σ

The underlying distribution is understood to be normal. The levels at which the tests are to be made are $-3.0, -2.5, -2.0, -1.5, -1.0, -.5, 0.0, .5, 1.0, 1.5, 2.0, 2.5$, and 3.0 measured in terms of standard deviations; the origin is the 50 per cent level. The first trial of each test is made at -2.5 . It is further assumed that the probability of an explosion at -3.0 is zero and at $+3.0$ is one.

b. Single Explosion plus m Trials Methods. First let us consider the Single Explosion plus m Trials Methods. The simplest case is to go up step by step to the first explosion and then report this level (m is zero). This method, commonly referred to as the Single Explosion Method, has been investigated for 1, 2, 3, 4, or 5 trials at a level. More generally a Single Explosion plus m Trials Method includes a set of trials (m) at a given level after the single explosion has occurred. Both the number of trials and the interval between the first explosion and the final level have been varied. Each such scheme can be identified by a triplet of numbers (k, m, h) representing the number of trials on each level in the up sequence, the number of trials at the final level and the number of intervals between the first explosion and the final level (positive h being measured in the direction opposite that of the up sequence).

In all but the $m = 0$ cases, the level reported depends on whether the final set of trials results in an explosion or not. An adjustment to the level of the first explosion is made in such a way that the expected value of the reported level for an explosion in the last set is the same as for non-explosions. Given the interval size and starting point (for the normal distribution) this adjustment minimizes the variance of the reported level.

One measure of the efficiency of a test is accuracy per trial. If each trial is equally expensive regardless of the result, then in repeated tests one obtains a specified degree of accuracy most cheaply by using the scheme with the maximum accuracy per trial. This quantity is given for a number of Single Explosion plus m Trials Schemes in Table 6.

For a given value of m and h , the accuracy per trial decreases with increasing k . That is, the greater the number of trials at each level on the up sequence, the more trials it takes to attain a given accuracy when other things remain fixed. For a given value of k and h , the variation with m is much smaller.

TABLE 6^{**}

Accuracy^{**} per Trial for
Single Explosion plus m Trials Methods
(Interval Size of .50')

k \ m h	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
1 0						(.321)						
1 1		.302	.331	.347	.353	.348	.334	.332	.299	.286		.277
1 2				.327	.350	.355	.341	.312	.283	.250		.236
1 3					.334	.347	.339	.308	.269	.242		.213
2 0						(.262)						
2 1	.260	.272	.281	.306	.291	.280	.266	.253	.245	.240	.236	
2 2			.271	.291	.296	.291	.270	.246	.232	.221	.216	
2 3			.259	.279	.303	.294	.272	.243	.220	.206	.199	
3 0						(.228)						
3 1		.245		.256	.258	.254	.248	.220	.216	.212	.210	.210
3 2				.259	.266	.253	.238	.217	.205	.196		
3 3				.260	.268	.257	.236	.213	.198	.187		
4 0						(.208)						
4 1						.216						
4 2						.226						
4 3						.231						
4 4						.230						
5 0						(.193)						
5 1						.201						
5 2						.205						
5 3						.211						
5 4						.215						
5 5						.218						
5 6						.215						

* For the estimated level in certain selected cases, see Tables 12 and 19 (pages 57 and 72).

** For Accuracy per Explosion, see Table 7 (page 49).

For Weighted Accuracy, see Table 9 (page 51).

() These values refer to the Single Explosion Method.

The maximum for fixed \underline{k} and \underline{h} is at a value of \underline{m} which is greater the greater \underline{k} is. In other words if a small number of trials is used at each level on the up sequence, a small number of trials should be used at the final level.

For given \underline{k} and \underline{m} , the maximum is near $h = 0$. In some cases it pays to jump up one or two levels and make the final trials. However, since at $k = 3$ the maxima for various \underline{m} are only slightly different from values for $h = 0$, it was thought that only the $h = 0$ cases need be studied for larger \underline{k} . In these cases there is a maximum for \underline{m} approximately equal to \underline{k} .

The maximum accuracy per trial for all Single Explosion plus m Trials Methods is given for $k = 1$, $m = 2$, $h = 0$, namely .355. In general, it takes an unnecessarily large number of trials if one takes more than one trial at a level on the up sequence.

In some sensitivity tests the expense of a trial resulting in a non-explosion is negligible compared to that of one ending in an explosion. In such instances one wishes to minimize the number of explosions in obtaining a given accuracy of estimate. The criterion for this is the accuracy per explosion. This criterion is tabulated in Table 7.

For the up sequence alone (i. e., Single Explosion Method) there can be only one explosion and then the criterion is simply the reciprocal of the variance. For these cases the variance decreases as the number of samples tested per level increases. In fact, it is easy to see that by increasing the number of trials indefinitely, one can be sure of ending on the same level that one starts on. However, when another set of trials is made at a final level, the expected number of explosions is greater than one. For a given pair, \underline{m} and \underline{h} , the accuracy per explosion increases with increasing \underline{k} .

For a given value of \underline{k} and \underline{h} accuracy per explosion may go up or down with increasing \underline{m} . For $h = 0$, the criterion decreases for $k = 1$ and $k = 2$, but it decreases with \underline{m} , having a relative maximum near $m = k$. For \underline{h} large, the criterion decreases and for \underline{h} small it increases.

For given values of \underline{k} and \underline{m} , the maximum of the criterion is at approximately $h = +2$. In such cases the final trials are made at a very low level where the probability of an explosion is small.

Of the schemes studied, the best from the point of view of number of explosions is the up sequence alone for $k = 5$ with a variance of .317. Of course, for higher \underline{k} the variance would be less. On the other hand with a high value of \underline{k} , one takes a considerably greater number of trials.

A third criterion which takes account of both number of trials and number of explosions is the weighted accuracy. This criterion assumes that the expense of a trial resulting in an explosion is 11 times as great as one resulting in a non-explosion. As pointed out in Section 4 it is a compromise between accuracy per trial and accuracy per explosion. The values for this criterion are given in Table 8.

In the case of the Single Explosion Method the maximum weighted accuracy is given by $k = 2, 3$ or 4 . For $k = 1$ the number of explosions is too large for the variance; for $k = 5$ or greater the number of trials is too large.

TABLE 8.

Weighted Accuracy^a for
Single Explosion plus m Trials Methods
(Interval Size of .5σ)

k \ m \ h	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
1 0						(1.14)						
1 1		.75	.86	.95	1.03	1.09	1.13	1.19	1.13	1.10		1.09
1 2				.87	.98	1.08	1.14	1.14	1.11	1.06		1.01
1 3					.93	1.05	1.14	1.15	1.10	1.05		.97
2 0						(1.22)						
2 1	.87	.93	1.00	1.15	1.17	1.20	1.21	1.20	1.19	1.18	1.17	
2 2			.93	1.05	1.14	1.21	1.22	1.19	1.17	1.13	1.12	
2 3			.88	.99	1.15	1.21	1.23	1.19	1.14	1.10	1.07	
3 0						(1.22)						
3 1		1.00		1.15	1.16	1.22	1.26	1.20	1.19	1.18	1.17	1.17
3 2				1.11	1.21	1.23	1.25	1.20	1.16	1.13		
3 3				1.09	1.19	1.24	1.23	1.19	1.14	1.10		
4 0						(1.22)						
4 1						1.21						
4 2						1.23						
4 3						1.24						
4 4						1.23						
5 0						(1.20)						
5 1						1.20						
5 2						1.20						
5 3						1.22						
5 4						1.23						
5 5						1.24						
5 6						1.22						

^a For Accuracy per Trial, see Table 6 (page 47).

For Accuracy per Explosion, see Table 7 (page 49).

() These values refer to the Single Explosion Method.

For the Single Explosion plus m Trials Method one has similar comparisons for fixed \underline{m} and \underline{h} . For example, for $m = 1$ and $h = 1$ the maximum is at $k = 3$. In fact, the larger \underline{m} is the greater is the \underline{k} for which the weighted accuracy is a maximum.

Varying \underline{m} for $h = 0$ and \underline{k} fixed yields a maximum for a value of \underline{m} depending on \underline{k} . The correspondence (i. e., for $h = 0$) is $k = 1, m = 0$; $k = 2, m = 0$; $k = 3, m = 3$; $k = 4, m = 3$; $k = 5, m = 5$. The maximum weighted accuracy for any \underline{k} and \underline{m} ($h = 0$) is 1.24.

If \underline{k} and \underline{m} are held fixed, the maximum weighted accuracy occurs near $h = 2$ for $k = 1$, near $h = 1$ for other \underline{k} . For large \underline{k} , however, the weighted accuracy for $h = 0$ differs little from this maximum.

Upon the basis of this investigation at interval size .50, five Single Explosion plus m Trials Methods were selected to be studied at different interval sizes. The results of this investigation are stated in the next section. The five methods chosen were:

1. $k = 1, m = 1, h = 0$
2. $k = 1, m = 1, h = 1$
3. $k = 1, m = 2, h = 0$
4. $k = 2, m = 1, h = 0$
5. $k = 2, m = 2, h = 0$.

c. Cascade Methods. A Cascade Method is a combination of two Single Explosion Designs, the first one starting at -2.5 and the second starting at a given number of intervals from the end level of the first run. The number of trials per level for each run need not be the same. Each Cascade Scheme can be identified with a triplet of numbers similar to those used for the Single Explosion plus m Trials Method. Here \underline{k} , \underline{m} and \underline{h} represent respectively the number of trials on each level in the first up sequence, the number of trials on each level in the second up sequence, and the number of intervals between the end of the first sequence and the start of the second one.

In all cases the level reported depends upon the end level of the first run, \underline{x} , and the end level of the second run, \underline{y} . An adjustment to the level of the first explosion is made so that the expected value of the reported level is the same for

all combinations of \underline{x} and \underline{y} for which $y-x$ has the same value. This adjustment minimizes the variance of the reported level.

The same measures of efficiency are used to compare the various Cascade Schemes as were used to compare the Single Explosion plus m Trials Schemes. Table 9 gives the values for accuracy per trial.

TABLE 9^{*}
Accuracy^{**} per Trial for
Cascade Methods

(Intervals of .5 σ)

k	m \ h	0	1	2	3	4
1	1	.383	.384	.372	.354	.332
	2	.365	.366	.353	.328	.298
	3		.331	.327	.300	.267
2	1				.297	
	2	.317	.313	.298	.276	.248
	3	.303	.302	.283	.259	
	4		.285	.267	.252	.221
3	1				.264	
	2	.285	.279	.266	.248	
	3		.264	.252	.230	

* For the estimated level in certain selected cases, see Table 15 (page 63).

** For Accuracy per Explosion, see Table 10 (page 55).

For Weighted Accuracy, see Table 11 (page 56).

Given values for \underline{m} and \underline{h} , the fewer trials one makes on a level in the first sequence, the greater will be the accuracy per trial. In every scheme tested the maximum accuracy per trial was for $k = 1$ for a fixed \underline{m} and \underline{h} .

Fixing \underline{k} and \underline{h} , one finds that changing the number of trials on a level in the second sequence does not cause as much variation as is obtained from changing \underline{k} . The maximum is at $m = 1$, even though \underline{k} increases. From this one concludes that the fewer trials one uses on a level in both sequences the greater is the efficiency.

For a given \underline{k} and \underline{m} , the maximum seems to be near $h = 0$. Actually it might result in slightly greater accuracy to continue up one or two levels before starting the second sequence. However, from the table it is evident that as the number of levels backed down decreases, the accuracy per trial increases at a decreasing rate. In the scheme $k = 1$, $m = 1$ or 2 the maximum is at $h = 1$, and the accuracy decreases if one starts the second sequence at the same level as the first sequence ends.

For all Cascade Schemes tested the maximum accuracy per trial results for $k = 1$, $m = 1$, $h = 1$. In general, the best results are for small \underline{k} , \underline{m} and \underline{h} . Changing any two of these quantities, the greatest accuracy per trial is still obtained for a small value of the third one.

Considering the accuracy per explosion, one finds it is merely a function of the variance. Since both runs end as soon as one explosion results, the number of explosions is always two, and the accuracy per explosion is the reciprocal of twice the variance. Table 10 gives this measure for a number of Cascade Schemes.

For a given pair, \underline{m} and \underline{h} , the more trials on a level of the first run the greater is the accuracy per explosion. Similarly for a given pair, \underline{k} and \underline{h} , the more trials on a level of the second run the greater is the accuracy per explosion. This is to be expected since the variance decreases (thus the reciprocal increases) as the trials per level increases. Increasing the number of trials per level in the second run has less influence than an increase in the number of trials per level in the first run.

Holding \underline{k} and \underline{m} constant one notes that for increasing \underline{h} the accuracy per explosion increases, but at a decreasing rate. It is quite obvious that it will level off, since a jump that makes the second run start below -2.5 has approximately the same effect as a smaller jump that makes the second run start at -2.5 . (The probability of an explosion below -2.5 is assumed to be zero.)

TABLE 10
Accuracy^a per Explosion for
Cascade Methods

(Intervals of .5σ')

k \ m \ h	0	1	2	3	4
1	1	1.45	1.56	1.65	1.71
	2	1.47	1.64	1.80	1.92
	3		1.66	1.86	2.02
2	1			2.03	
	2	1.90	2.06	2.18	2.26
	3	1.92	2.14	2.29	2.42
	4		2.16	2.36	2.51
3	1			2.23	
	2	2.16	2.31	2.41	2.47
	3		2.40	2.54	2.62

^a For Accuracy per Trial, see Table 9 (page 53).

For Weighted Accuracy, see Table 11 (page 56).

Of the schemes tested, the best from the point of view of accuracy per explosion is the one with $k = 2$, $m = 4$, $h = 4$. However, larger k , m and h , if studied, would undoubtedly have given greater values for accuracy per explosion.

The values for the weighted accuracy are given in Table 11. Fixing m and h , the weighted accuracy is greater at $k = 3$ for small values of h and greater at $k = 2$ for larger values of h . For fixed k and h weighted accuracy varies little with changes in m . The weighted accuracy is slightly larger at $m = 2$ for $k = 1$, at $m = 3$ for $k = 2$ and at either $m = 2$ or 3 for $k = 3$. Fixing k and m gives the maximum weighted accuracy at $h = 2$ or 3 , though a change in h has little effect.

The maximum weighted accuracy for all of the schemes tested was 1.27 for $k = 2$, $m = 3$, $h = 2$. Judging from the quantities in Table 11, there is only a slight advantage in any one method over any other since the values listed vary only from 1.05 to 1.27.

TABLE 11
 Weighted Accuracy^{*} for
 Cascade Methods
 (Intervals of .5σ)

k \ m h	0	1	2	3	4
1	1	1.05	1.11	1.14	1.15
	2	1.05	1.13	1.19	1.21
	3		1.11	1.19	1.21
2	1			1.21	
	2	1.20	1.24	1.26	1.24
	3	1.17	1.25	1.27	1.25
	4		1.23	1.25	1.23
3	1			1.21	
	2	1.23	1.26	1.26	1.24
	3		1.26	1.26	1.23

* For Accuracy per Trial, see Table 9 (page 53).

For Accuracy per Explosion, see Table 10 (page 55).

Upon the basis of our investigation at interval size .5σ, we have selected five Cascade Methods for closer study in the next section. In particular, we shall be concerned with their behavior when the interval size is changed. The methods chosen are as follows:

1. $k = 1, m = 1, h = 1$
2. $k = 1, m = 1, h = 3$
3. $k = 1, m = 2, h = 3$
4. $k = 2, m = 1, h = 3$
5. $k = 2, m = 2, h = 1.$

10. Staircase Methods at Different Step Sizes.

a. General. For most staircase methods the size of the step has an effect on the outcome of the test. However, for some methods such as the Up and Down (for step sizes less than $1.5\sigma'$) the step size affects mainly the accuracy of the final estimate, but not its average value. On the other hand the average outcome of a method like the Single Explosion Method depends to a considerable extent on the size of the step. In this case, if the step size is extremely small the percent point estimated is relatively small because when many trials are made where the probability of an explosion is small eventually an explosion occurs (i. e., one moves up very slowly). If the step size is large, the test quickly arrives at a higher per cent point. Figure 9 indicates the average level at which the explosion occurs for different step sizes. A table of the expected levels is given below.

TABLE 12

Average Level Estimated by Single Explosion Design

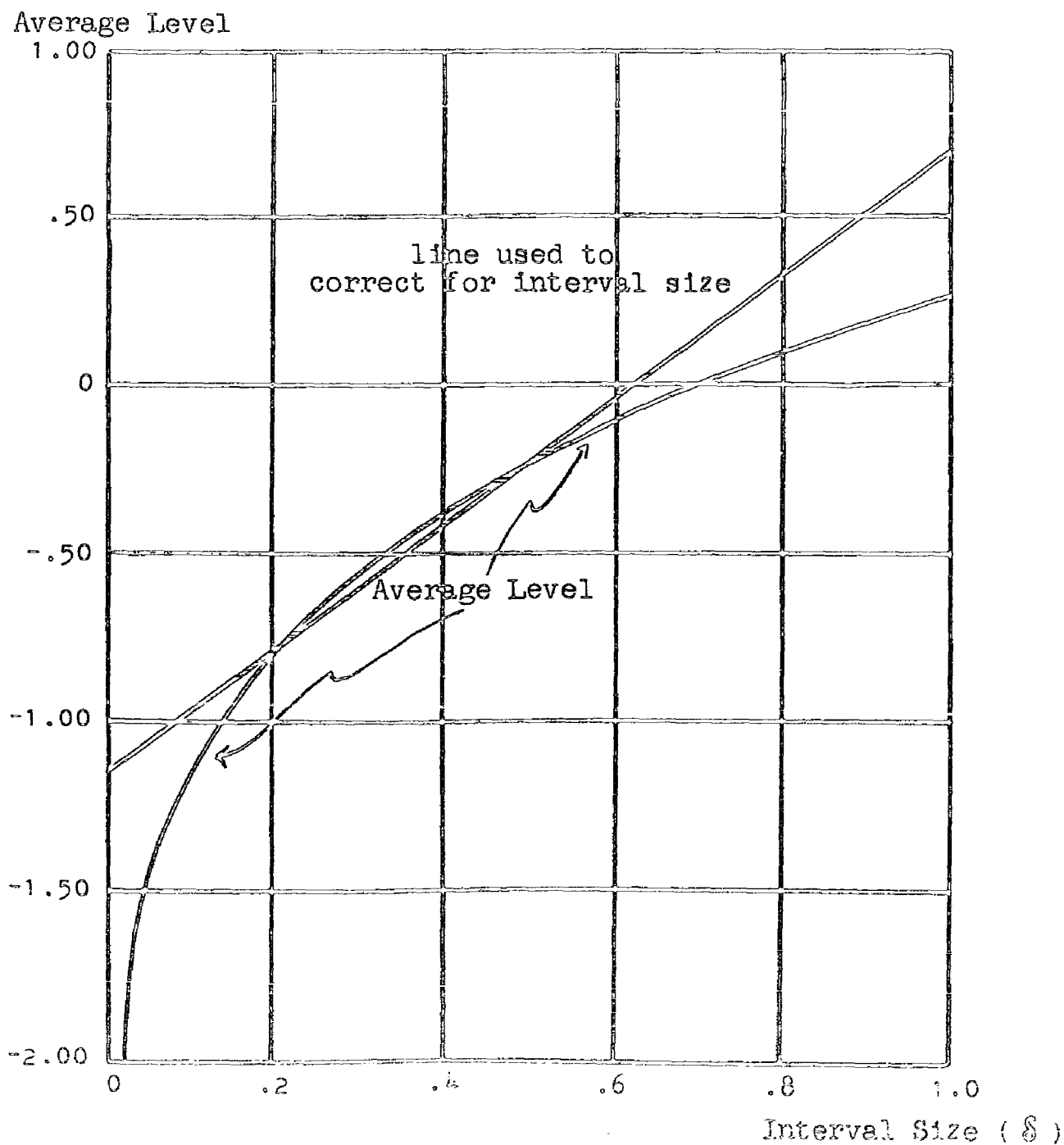
(one trial on a level)

Step Size	Average Level
.025	-1.711
.05	-1.444
.1	-1.140
.2	-.790
.5	-.241
.7	-.005
1.0	+.274

b. Correction for step size. It would be desirable to use a design for which the average estimated point does not depend on the step size. Unfortunately, most of the efficient designs have this step size effect and it is therefore necessary to include in the analysis accompanying a design a compensating correction. One cannot completely eliminate the effect but if the approximate step size is known, it can be reduced to an almost negligible amount. The general approach to be used in this problem has been outlined in Section 8.

Figure 9

Average Level Estimated by the
Single Explosion Design
(one trial on a level)



An example will show the kind of correction that is feasible. The Single Explosion Method estimated $-.790\sigma'$ for a step size of $.2\sigma'$ and $-.241\sigma'$ for a step size of $.5\sigma'$. We desire an adjustment, which is independent of the step size, such that the adjusted average level will be the same for the two step sizes. In other words, we wish to determine a constant, c , such that

$$-.790\sigma' - (.2\sigma') c = b\sigma', \text{ and}$$

$$-.241\sigma' - (.5\sigma') c = b\sigma'.$$

From these equations c is immediately found to be equal to

$$\frac{-.790 - .241}{-.2 - .5} = 1.83.$$

This procedure can be summarized as follows:

Step Size	Average first explosion at	Average level of first explosion minus 1.83 step sizes
$.2\sigma'$	$-.790\sigma'$	$-1.156\sigma'$
$.5\sigma'$	$-.241\sigma'$	$-1.156\sigma'$

In Figure 9 this correction can be illustrated by drawing a straight line through the points on the curve at $.2$ and $.5$. The y-intercept of the line is approximately -1.156 and the slope is approximately 1.83 .

To correct for any other two step sizes we draw the corresponding line. If the curve were a straight line then the same line would correct for all points; that is, a constant times step size subtracted from the estimate would give the same average irrespective of step size. Small curvature is one of the desirable features of a good test.

For each pair of step sizes, one can make an exact correction in this manner. If this difference between the two step sizes is small, the correction will hold good approximately for the region between. In Section 5 some rough corrections are suggested. For example, between $.5$ and 1.0 step sizes it is suggested that one

subtract 1 step. This correction gives

Interval Size	Average Level
.5	-.741
.7	-.705
1.0	-.726

All of these values are near $-.67$ which is the 25 per cent point.

c. The NPF and NPMI Methods. The NPMI method (i. e., NPF Inverted) is simply the reverse of the NPF Method. In the former case we estimate the 10 per cent point (approximately) and in the latter, the 90 per cent point. For the sake of convenience, we shall write in detail only of the NPMI design. It is understood that analogous statements hold true for the NPF Method.

Below is given a table which indicates the expected level and the variance of the reported level for various step sizes.

TABLE 13

Characteristics of the NPMI Method

Step Size	Expected Level	Variance	Average No. of Trials	Accuracy per Trial	Average No. of explosions	Accuracy per explosion
.1	-1.300	.2250	19.7	.226	1.65	2.69
.2	-1.153	.2497	15.0	.267	1.87	2.14
.5	-1.219	.3180	10.2	.308	1.96	1.60
1.0	-1.494	.4152	8.1	.297	1.77	1.36

For this method the level estimated depends on the step size in such a way that for either very small or very large step sizes the per cent point is very small while for intermediate step sizes, the per cent point may be as large as 13 per cent. In Section 5 instructions are given for correction factors for various pairs of step sizes. For a step size of about .3 no correction is needed. The average number of trials increases as the step size decreases. On the other hand the variance decreases. That is, for large step size one makes few trials but moves

very rapidly. Thus there is considerable variability. The accuracy per trial has a broad maximum near a step size of $.6\sigma$.

For a particular interval size one reduces the variance by using the information given by knowledge of where the first explosion occurs. This adjustment is made as suggested in Section 8. This means that for each value of the difference between final level and first explosion level one calculates the expected level and adjusts for the differences of these levels. However, the reduction in variance in this instance is fairly small, less than 10 per cent. The expected levels given in Table 13 are uncorrected, while the variance is with the correction applied.

The average number of explosions also depends on the interval size. The maximum is at about an interval size of $.5$. Since the variance is smallest at the smallest interval size, the accuracy per explosion is largest at small interval sizes. Hence, if one wishes to keep the number of explosions down, one uses the smallest step size feasible. This will also maximize the weighted accuracy.

d. Cascade Methods. As indicated in Section 9 five Cascade Methods were selected for a study of the effect of step size on expected level and efficiency. The characteristics of each method were computed for step sizes of $.2$, $.5$, and 1.0 . To make certain that at step size 1.0 (where the effect of the relation of intervals to the origin is greatest) testing on the levels -2.5 , -1.5 , $-.5$, etc. instead of -3.0 , -2.0 , -1.0 , etc. made no difference, starting at -3.0 was compared with starting at -2.5 for each method. Essentially the only difference was that starting at -3.0 added one half the number of trials per level to the average number of trials.

In a test with a given Cascade Method two numbers result, namely, the levels of the first and second explosion. Another way of recording this information is to take the level of the first explosion, and the level of the second explosion minus the level of the first explosion. This difference has been designated as \underline{r} . In Figure 10 we have graphed, for a particular method and step size, the expected level of the first explosion for each value of \underline{r} . This information is for the $k = 1$, $m = 1$, $h = 1$ scheme (step size of $.2\sigma$). If, in making a test, one were certain that the step size were $.2$, then the most efficient estimate would be to add to the level of the first explosion an amount depending on \underline{r} so that whatever the value of \underline{r} one

Figure 10

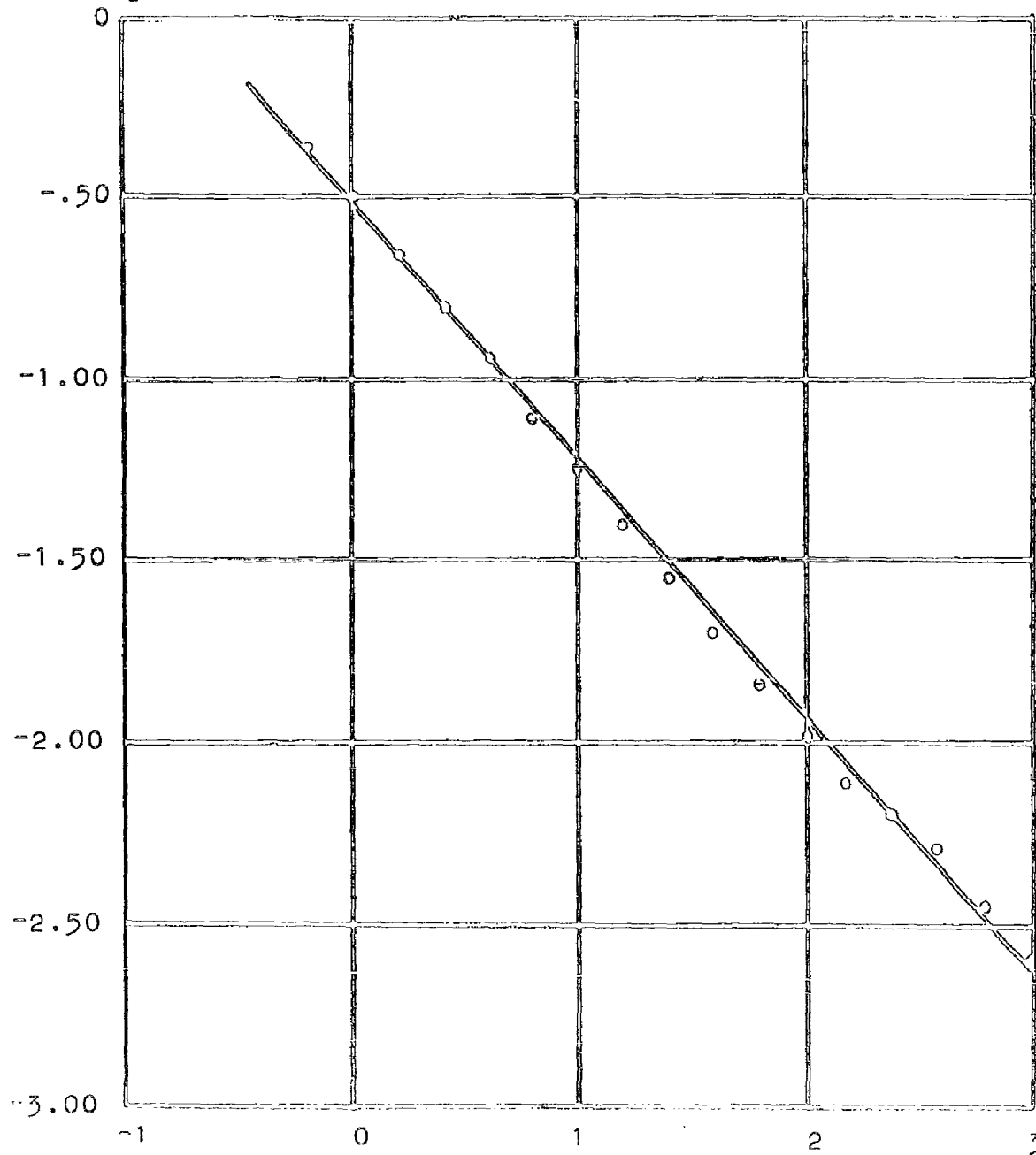
The Expected Level of the First Explosion as
a Function of r for a

Cascade Design

($k = 1, m = 1, h = 1$)

Interval size of $.2 \sigma$

Expected Level
of First Explosion



$$J - X = r$$

obtains the same level on the average. This correction minimizes the variance.

The effect of r is nearly linear. Hence, we can simply add to the level of the first explosion an amount which is the product of r and a certain number, a . In the case of .2 step size the proper number a is .74. Then one estimates -.50 (i. e., the 31 per cent point). The variance is increased negligibly (by .0001) by using a linear correction instead of the exact correction. In Table 14 are given the values of the correction factor a for four step sizes, and also the average reported level.

TABLE 14
Correction Factors for a Cascade Design
($k = 1, m = 1, h = 1$)

Step Size	Correction Factor a	Corrected Average Level
.1	.78	-.83
.2	.74	-.50
.5	.66	-.03
1.0	.58	+.38

If the correction factor were the same for all step sizes, one could set up a single adjustment for use at any step size. Then one would merely need to correct for variations in the average level as in the NPFI scheme. Since the correction factors do vary, we must use some compromise. For example, suppose we believe that the step size is between .2 and .5. If we add $2/3$ of $(y-x)$ to the first end point we would get almost the same effect as if we used .74 (for .2) or .66 (for .5). This is, of course, the same as averaging the two end points and weighting the second twice as heavily as the first. Then to correct for the average level we shall subtract 1.7 interval sizes (see discussion of Single Explosion Design in (b) of this section). Such a procedure estimates the 12 per cent point. At .2 there is a bias of .010 and a contribution to the variance of .0016; at .5 there is neither bias nor contribution to the variance.

Table 15 gives the correction factors for $(y-x)$ and the average levels for other Cascade Schemes. Adjustments similar to the one considered above can be made for each procedure.

TABLE 15
Values of the Correction Factor and the Average
Reported Level for Various Cascade Designs

Interval Size	k=1, m=1, h=3		k=1, m=2, h=3		k=2, m=1, h=3		k=2, m=2, h=1	
	<u>a</u>	Average Reported Level	<u>a</u>	Average Reported Level	<u>a</u>	Average Reported Level	<u>a</u>	Average Reported Level
.2	.65	-.64	.81	-.81	.51	-.90	.73	-.90
.5	.54	-.20	.64	-.44	.45	-.46	.64	-.50
1.0	.50	+.28	.57	+.01	.44	+.01	.57	-.13

In general, for increasing step size we have increasing variance but decreasing average number of trials. The net result is that the accuracy per trial increases to a point and then levels off or decreases. In several cases the maximum accuracy per trial occurs near a step size of .5.

It is characteristic of all of these methods that the smaller the step size the lower the per cent point estimated. Suppose that we plot the estimated point against step size and draw the tangents to the curve at the known points. Call the intercepts on the vertical axis M_{δ} . These represent the average levels when the correction is ideal for step sizes near δ . M_{δ} also increases with δ .

Figure 11 is the plot of M_{δ} against accuracy per trial for different interval sizes and different Cascade Methods. Up to interval sizes of about .5 the points lie quite well on one curve, which means that accuracy per trial increases as M increases, regardless of the test. This implies that it is more expensive by any test to estimate extreme per cent points than (relatively) moderate ones. Furthermore, it does not make too much difference which of these tests one uses. The values of M_{δ} and accuracy per trial for the Cascade Methods are given in Table 16.

Figure 11

Plot Showing Relationship between Accuracy per Trial and M_S
for Various Cascade Designs

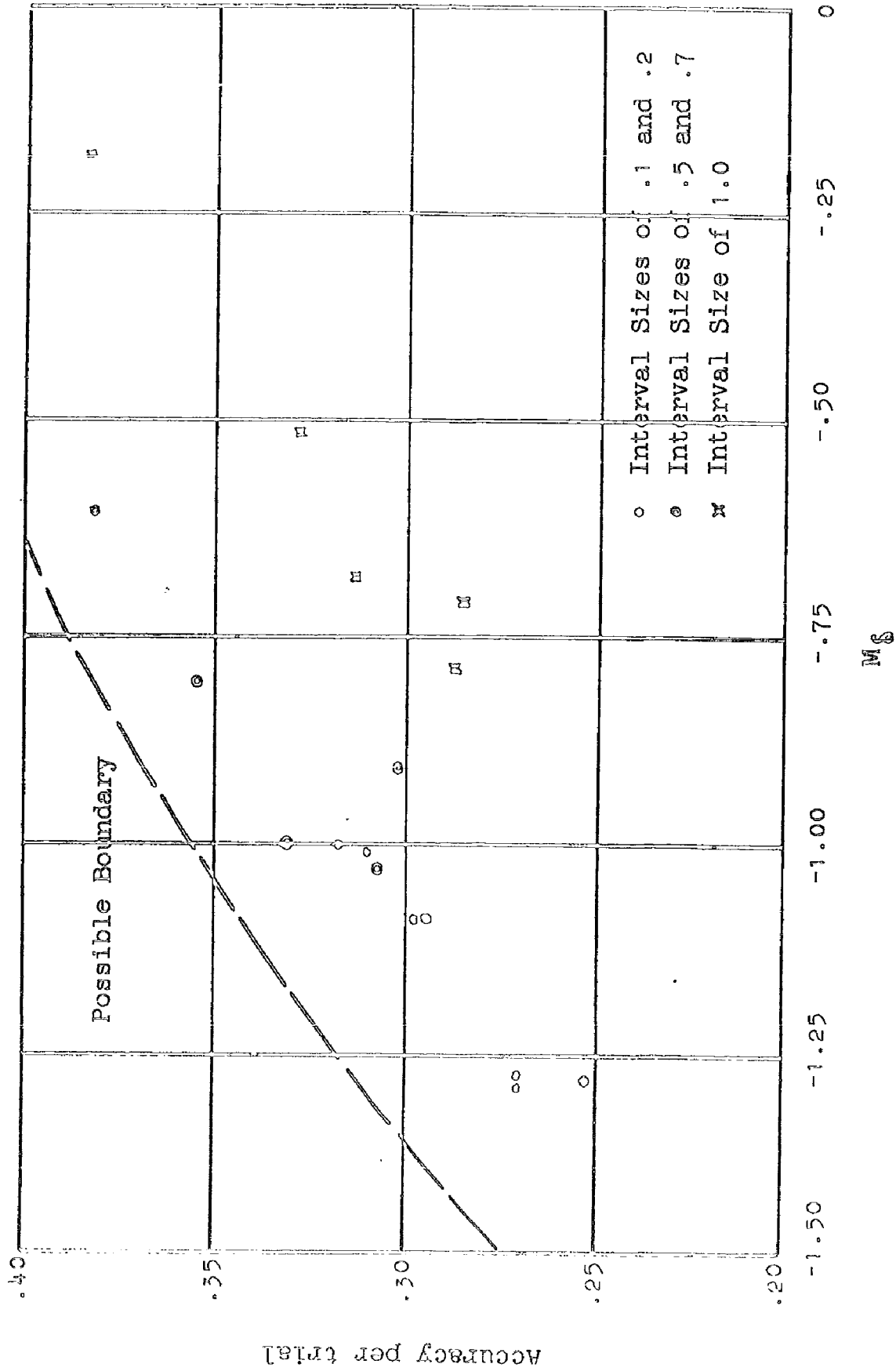


TABLE 16

Values of M_δ and Accuracy per Trial
for Cascade Methods

Interval Size	k=1, m=1, h=1		k=1, m=1, h=3		k=1, m=2, h=3		k=2, m=1, h=3		k=2, m=2, h=1	
	M_δ	Acc./ Trial	M_δ	Acc./ Trial	M_δ	Acc./ Trial	M_δ	Acc./ Trial	M_δ	Acc./ Trial
.1	-1.28	.269								
.2	-1.00	.318	-1.01	.310	-1.07	.294	-1.29	.251	-1.28	.269
.5	-.61	.384	-.80	.354	-1.00	.328	-1.07	.297	-1.03	.313
.7					-.91	.303				
1.0	-.17	.387	-.51	.327	-.77	.284	-.72	.280	-.67	.316

The number of trials on each level in the successive runs of a Cascade Method should clearly affect the accuracy per explosion. Empirically the product of (i) the average over the runs of these numbers, and (ii) the number of levels per σ is the important quantity. We denote this product as trials per σ . For example, the scheme involving one trial per level moving up twice provides 1 trial per σ when the step size is 1, 2 trials per σ when the step size is .5, etc. The scheme which proceeds up by one trial to an explosion and then by two trials per level to the second explosion provides 1.5 trials per σ when the step size is 1.0.

Figure 12 indicates this relationship for the 5 Cascade Methods. It is clear that the points are nearly on a single curve. To raise the accuracy per explosion one must increase, in any way, the number of trials per σ . This can be done by decreasing the interval size or by increasing the number of tests per level, both of which increase the average number of trials, or by changing to a method with fewer runs and more trials per level. This latter need not increase the average number of trials. Apparently the practical maximum accuracy per explosion is between 4 and 6. The values for trials per σ and accuracy per explosion for the Cascade Methods are given in Table 17.

Figure 12

Plot Showing Relationship between Accuracy per Explosion
and Trials per σ for Various Cascade Designs

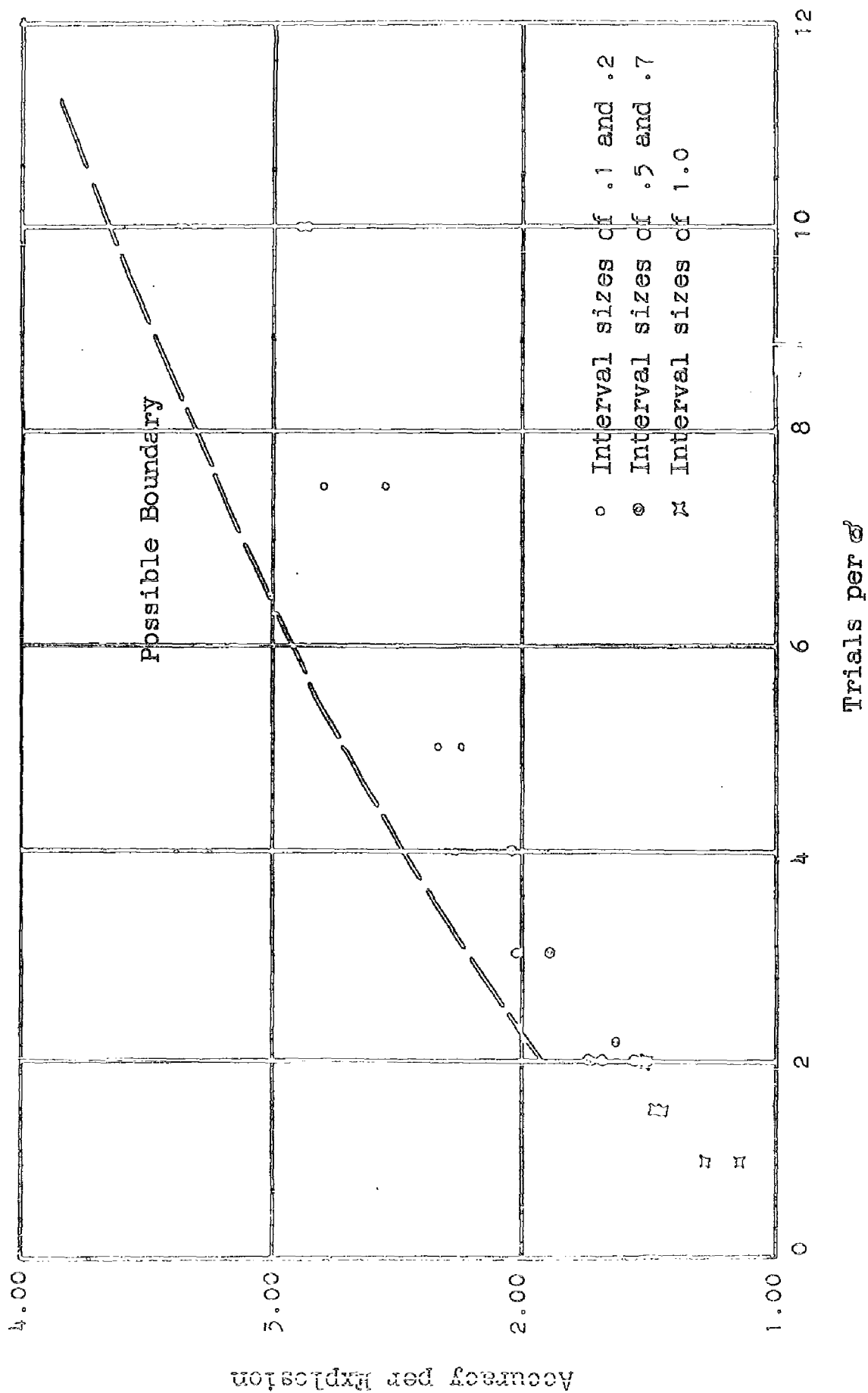


TABLE 17

Values of Trials per σ' and Accuracy per Explosion
for the Cascade Methods

Interval Size	k=1, m=1, h=1		k=1, m=1, h=3		k=1, m=2, h=3		k=2, m=1, h=3		k=2, m=2, h=1	
	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.
.1	10	2.89								
.2	5	2.25	5	2.56	7.5	2.53	7.5	2.81	10	2.69
.5	2	1.56	2	1.71	3	1.92	3	2.03	4	2.06
.7					2.14	1.64				
1.0	1	1.15	1	1.27	1.5	1.44	1.5	1.46	2	1.52

e. Single Explosion plus m Trials Methods. These methods can be treated in a fashion similar to that of the Cascade Methods. On a test with a given method two pieces of data result, namely, the level of the first explosion and whether an explosion occurs in the subsequent trial or set of trials. The reported level is the level of the first explosion modified suitably according to the result of the next trial (or trials). Table 18 gives the expected level of the first explosion for an explosion or for a non-explosion at different interval sizes for one design.

TABLE 18

Single Explosion plus m Trials
(k = 1, m = 1, h = 0)

Step Size	Expected Level for Explosion	Expected Level for Non-Explosion	Difference in Expected Levels	Variance
.2	-.39	-.92	.53	.333
.5	-.17	-.55	.72	.44
1.0	-.67	-.29	.96	.569

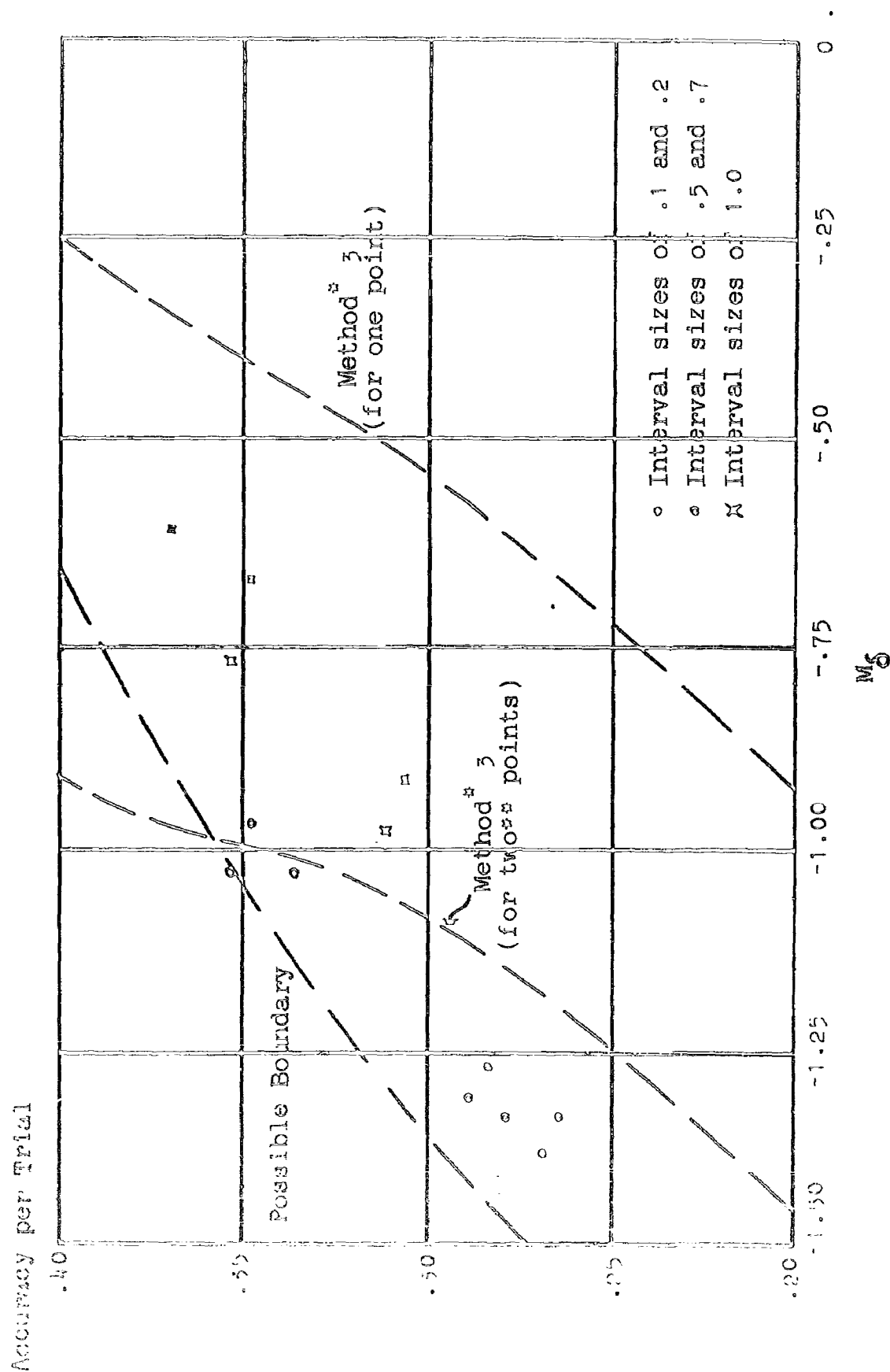
The type of adjustment possible for minimizing the variance of the reported level with the Single Explosion plus m Trials Designs has been discussed in Section 8. Following this example, we see that at step size .2 one should subtract 2.6 ($= .53/.2$) step sizes when an explosion occurs on the last trial to match the $-.92$ average for a non-explosion on the last trial with the minimum variance (.333). At step size .5, however, the matching adjustment is 1.4 ($= .72/.5$) step sizes to estimate $-.55$. Suppose we use the correction of 1.4 when the step size is .2. The expected level is now $-.85$ (instead of $-.92$) and the variance is .345 (instead of .333). Then we should subtract one step size (see Section 10(b)) to compensate for effect of step size on average level. This adjustment gives an estimated level of -1.05 ($= -.85 -.2 = .55 -.5$). It is clear that similar adjustments can be made for any Single Explosion plus m Trials Design and any pair of step sizes. Of course, sometimes the resulting variance is far from the minimum for a given step size.

Similar characteristics, i. e., expected level for explosion, expected level for non-explosion(s) and variance, for the other Single Explosion plus m Trials Designs are given in Table 19.

Let us now see how efficient these tests are in terms of accuracy per trial. In Figure 13 are plotted the accuracies per trial against the levels estimated by ideal correction (for interval size). The curve that these points approximate is roughly the same as the corresponding one for the Cascade Methods. In other words, the accuracy per trial depends mainly on the level estimated, not on the particular test. Values for M_g and accuracy per trial are given in Table 20.

The analysis of accuracy per explosion reveals that the Single Explosion plus m Trials Designs are in this respect, too, similar to the Cascade Methods. Figure 14 shows the relationship between accuracy per explosion and "trials per σ " (see preceding discussion of Cascade Methods). In this instance we neglect the additional m trials in determining "trials per σ ". The points lie approximately on the same curve as we obtained for the Cascade Methods. The implication is that if one increases the number of trials per level and decreases the step size he will increase the efficiency in terms of explosions to achieve the desired accuracy. Values of trials per σ and accuracy per explosion are given in Table 21.

Figure 13
 Plot Showing Relationship Between Accuracy per Trial
 M_S for Various Single Explosion plus m Trial Designs



* See Table 1, page 10

** See discussion on page 95.

Figure 14

Plot Showing Relationship Between Accuracy per Explosion
and Trials per ϕ' for various Single Explosion
plus m Trials Designs

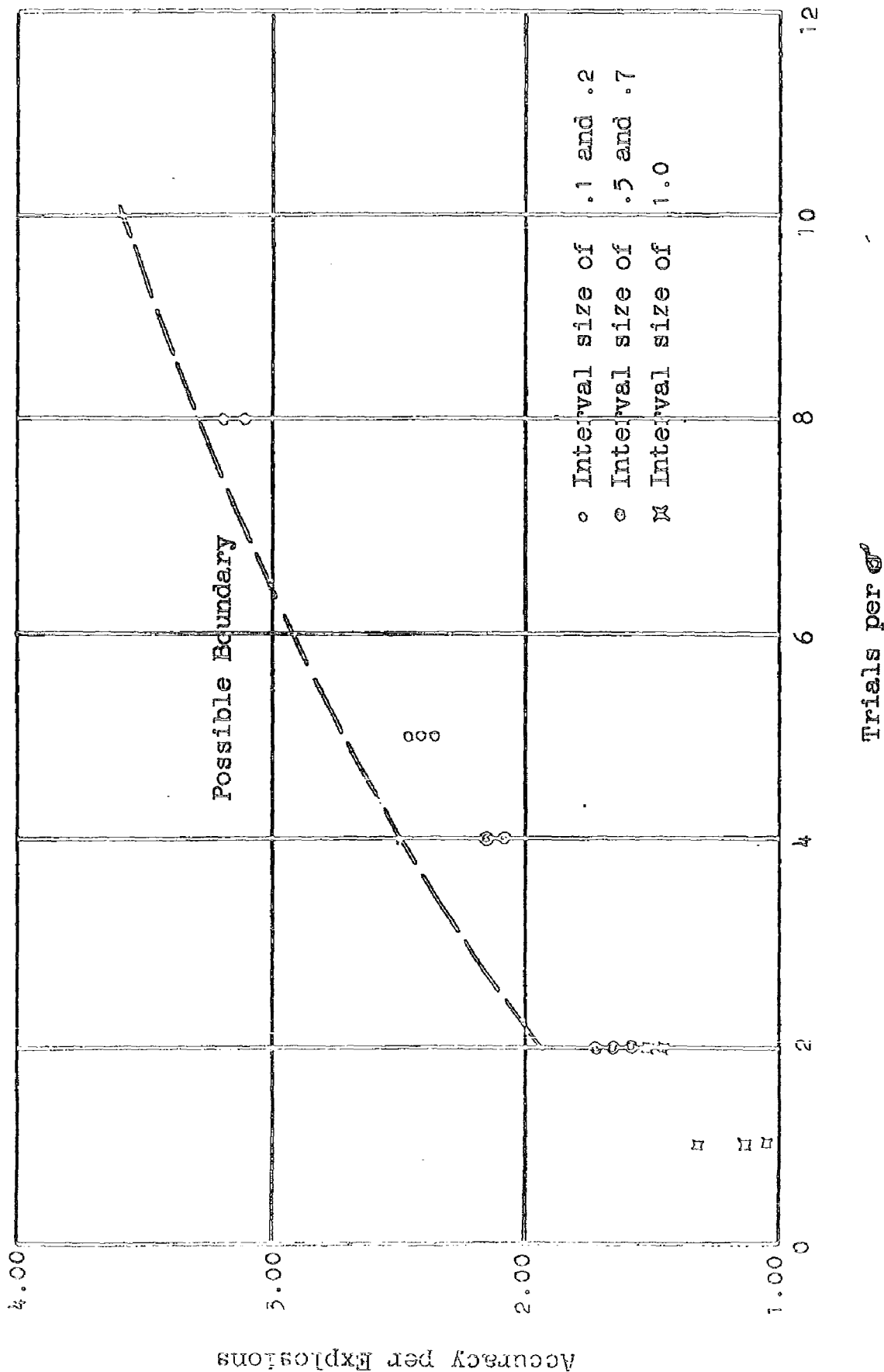


TABLE 19

Characteristics of the Single Explosion
plus m Trials Methods

Step Size	k = 1, m = 1, h = 1			k = 1, m = 2, h = 0		
	Expected Level for Explosion	Expected Level for Non-Explosion	Variance	Expected Level for Explosion	Expected Level for Non-Explosion	Variance
.2	-.36	-.90	.3400	-.44	-1.03	.2980
.5	+.29	-.45	.4580	+.08	-.76	.3970
1.0	+.95	-.01	.6033	+.56	-.61	.5441

Step Size	k = 2, m = 1, h = 0			k = 2, m = 2, h = 0		
	Expected Level for Explosion	Expected Level for Non-Explosion	Variance	Expected Level for Explosion	Expected Level for Non-Explosion	Variance
.2	-.76	-1.21	.2726	-.79	-1.27	.2532
.5	-.24	-.83	.3650	-.31	-.96	.3280
1.0	+.23	-.54	.4709	+.13	-.77	.4276

TABLE 20

Values of M_δ and Accuracy per Trial for Single
Explosion plus m Trials Methods

Interval Size	k=1, m=1, h=0		k=1, m=1, h=1		k=1, m=2, h=0		k=2, m=1, h=0		k=2, m=2, h=0	
	M_δ	Acc./ Trial	M_δ	Acc./ Trial	M_δ	Acc./ Trial	M_δ	Acc./ Trial	M_δ	Acc./ Trial
.2	-1.37	.272	-1.32	.266	-1.25	.285	-1.68	.215	-1.63	.221
.5	-.97	.348	-1.06	.335	-1.06	.355	-1.27	.280	-1.31	.291
1.0	-.62	.368	-.71	.347	-.76	.354	-.90	.306	-.98	.312

TABLE 21

Values of Trials per σ' and Accuracy
per Explosion for Single Explosion
plus m Trials Methods

Interval Size	k=1, m=1, h=0		k=1, m=1, h=1		k=1, m=2, h=0		k=2, m=1, h=0		k=2, m=2, h=0	
	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.	Trials per σ'	Acc./ Exp.
.2	5	2.40	5	2.45	5	2.37	10	3.17	10	3.09
.5	2	1.59	2	1.71	2	1.56	4	2.11	4	2.08
1.0	1	1.11	1	1.28	1	1.05	2	1.47	2	1.44

f. Estimating the step size. Since the level estimated depends on the step size it is desirable to have some practical way of estimating the step size when it is unknown. This can be done when several tests are made on the same sample of explosive under the same conditions. The general idea is that we use the set of trials to estimate the variation of the reported level in terms of step sizes and thereby estimate the step size.

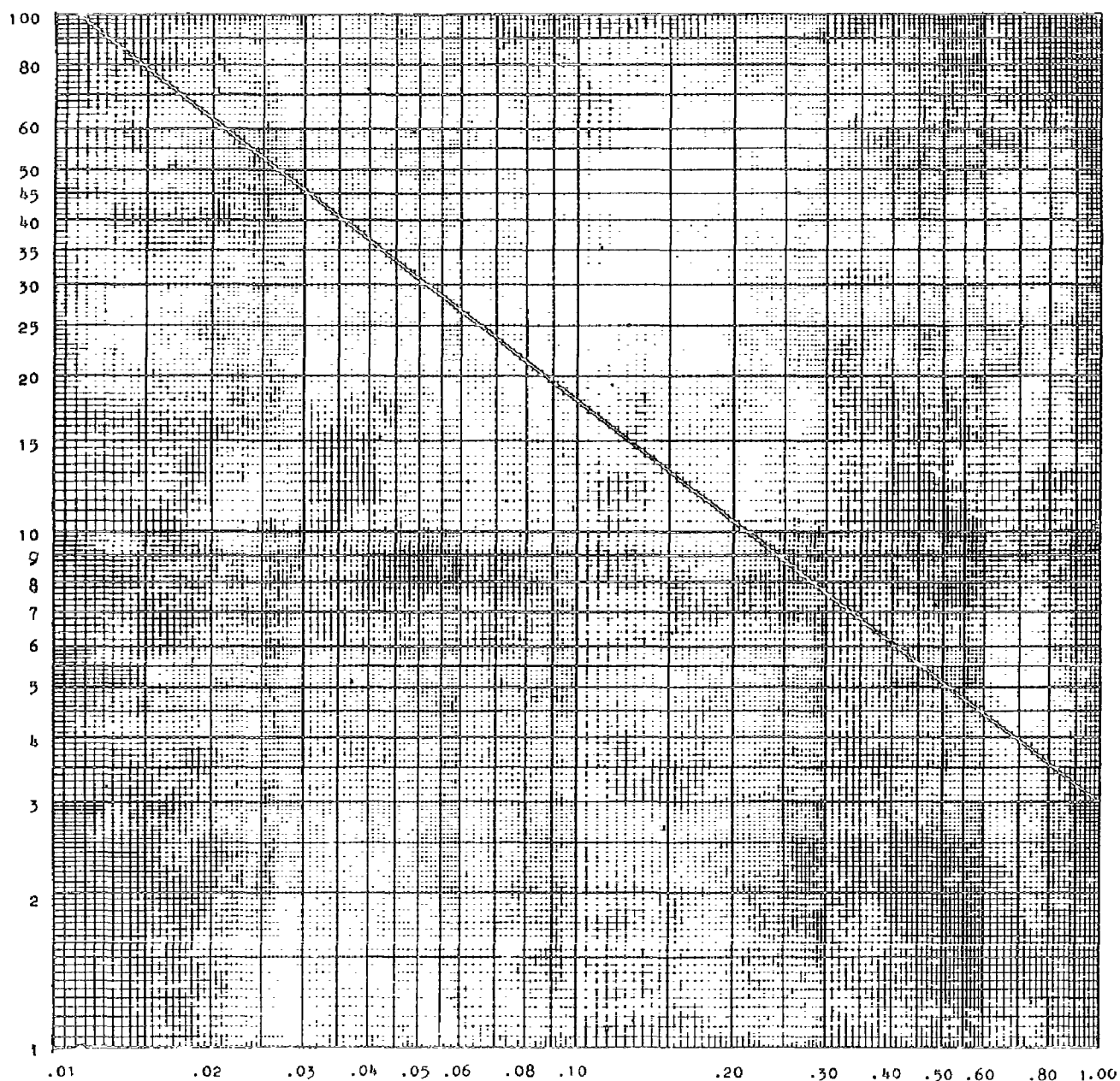
Suppose we have a set of 5 tests. Let us use the sum of deviations about the median as our measure of variation. This quantity, which we term the Total Deviation is simply the largest level less the smallest added to the difference of the next largest and next smallest. For the Single Explosion Scheme Figure 15 gives the expected value of the total deviation (expressed in step sizes) for different step sizes.^{*} If we measure the total deviation in terms of steps, it will always be an integer because each explosion occurs on one of the given levels. Using the graph in reverse we can estimate the step size from the observed total deviation. If more than 5 tests are made (say N tests), compute the total deviation by taking the sum of differences of largest and smallest observed levels, etc. Multiply this total deviation, expressed in step sizes, by $\frac{9}{2N-1}$ to obtain a quantity whose expected value is approximately the same as the total deviation (in step sizes) in a set of 5 tests.

* When the underlying distribution is nearly normal, the total deviation (for a set of 5 cases) averages 3.3 times the standard deviation.

Figure 15

Expected Value of Total Deviation (about Median) in Interval Sizes
for 5 Single Explosion (one trial on each level) Tests

Total Deviation
(in Interval Sizes)



Interval Size (in σ)

On the basis of our estimate of the step size we can then correct the average level of the 5 tests so as to estimate the 10 per cent point (or any other point). Figure 16 shows what the correction to be subtracted should be, both in terms of δ and in terms of step sizes, for the Single Explosion Design.

11. The Sequential Method.

a. General discussion. In Section 3 of this report a "staircase" method has been defined as a method where the severity of the next trial or group of trials is directly determined by the last trial or group of trials. For each method which has been discussed up to this point a rule of procedure has been given which, when carried through to the completion of the test, determined the characteristics of the estimated percentage point. These rules of procedure have been chosen because they furnish an estimated percentage point in an efficient manner, efficiency referring to the criteria of accuracy per trial, accuracy per explosion, or weighted accuracy. No particular attention has been paid to what the method accomplishes at a fixed level.

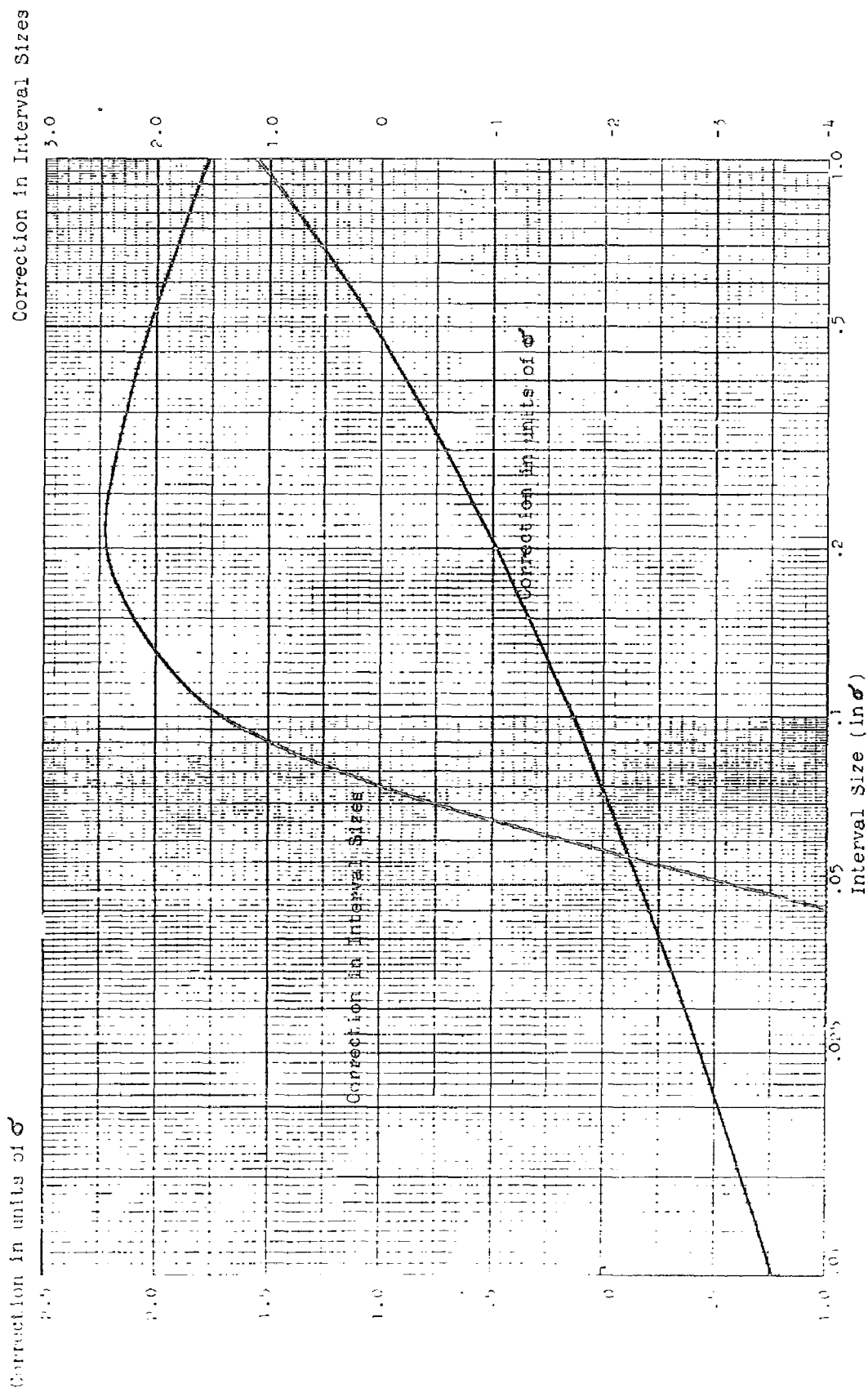
A systematic, rather than empirical, approach to the problem can be obtained by focusing attention primarily upon the relation between the results of testing at a given level and the percentage point which is to be estimated. Clearly the level corresponding to the desired percentage point must either (1) be above the level at which the testing is taking place, (2) lie on the test level, or (3) lie below the test level. Now if a reasonable criterion can be obtained which will distinguish between these three possibilities upon the basis of trials made at the level, then testing at successive levels will give directly usable evidence concerning the desired percentage point. For suppose that at level x , the criterion indicates that the desired level is above level x . Then if testing is done at level $x + \delta$ and the criterion indicates that the desired level is below this level, there is evidence that the desired level is between levels x and $x + \delta$.

There are certain rather obvious ways in which such a criterion can be arrived at. For example, one might simply carry out ten (or any other fixed number of) trials on a level and calculate the per cent explosions. If this experimental percentage were lower than the desired percentage, the next ten trials would be conducted on the next higher level. If it were lower, the next ten trials would be

Figure 16

Correction to be Subtracted from Average Level of First Explosion
in order to Estimate the 10 Per cent Point

Single Explosion Design, one trial on each level



conducted on the next lower level. This procedure would then be continued until the first time that the results on one level indicated that the desired percentage point was above this level, and the results on the next higher level indicated that the desired percentage point was below this level. For the moment the actual level assigned to the desired percentage will be neglected, although it would necessarily be taken between the two final levels. It should be noticed that this procedure is related to the standardized Picatinny Method.

Once the number of trials to be used on a level has been determined, the probability can be calculated that any particular decision will be obtained on a specified level. Thus, if the number of trials made on a level is ten, the probability of an explosion on a level is p_x and the 15 per cent point is to be estimated, then the probability is $(1-p_x)^{10} + 10(1-p_x)^9 p_x$ that the criterion tells us to go up a level after 10 trials on level x . Similarly, the probability that the criterion tells us to go down one level is $1 - (1-p_x)^{10} - 10(1-p_x)^9 p_x$. In this example there is no provision made for saying that the level tested corresponds to the 15 per cent point, and so the probability of this decision is zero. The probability that the decision will be to move up one level is graphed as a function of p_x in Figure 17. This curve is ordinarily called the Operating Characteristic Curve or the Operating Characteristic (abbreviated OC Curve or OC). Once the OC for a criterion requiring a fixed number of trials at a level has been determined it is possible to compute the average number of trials required to complete one determination of a desired percentage point, and also the variance of this estimate.

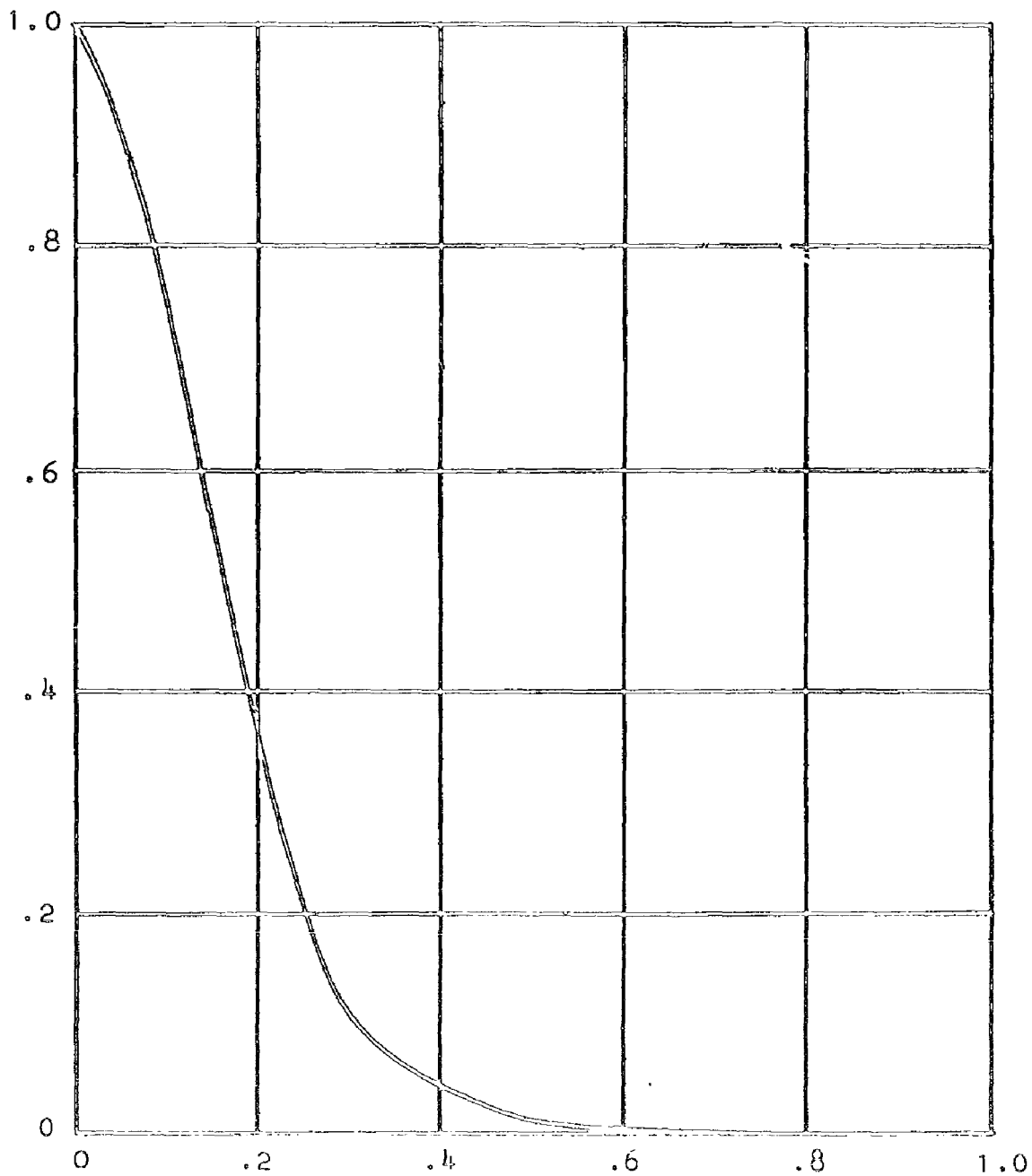
It is apparent from the OC given in Figure 17 that at a single level one may commit one of two errors. First, if the testing is being done at a level where p_x is less than .15, there is a probability of moving down one level when one should be moving up a level, and second, if testing is being done at a level where p_x is greater than .15, there is a probability of moving up one level when one should be moving down a level. These errors would both be zero if the OC had the value one for all values of p_x less than .15 and zero for all values of p_x greater than .15. However, such a curve could only be produced by setting up a plan with an infinite number of trials. Accordingly it is customary to describe such a plan as this by choosing a value p_1 less than .15 and a value p_2 greater than .15, and specifying the probability that the first type of error will be made if $p_x = p_1$ and the probability that the second type of error will be made if $p_x = p_2$.

Figure 17

Probability of Moving to Level $x+\delta$ After Making
Ten Trials on Level x

(see text for description of test procedure)

Probability



Probability of an Explosion on Level x , p_x

Thus, if testing is being done at a level where $p_x = p_1$, the probability of an error is designated by α , and if testing is being done at a level where $p_x = p_2$, the probability of an error is designated by β . For $p_x < p_1$ the probability of an error is less than α , and for $p_x > p_2$ the probability of an error is less than β . It should be noted that in this particular, simple example (for given p_1 and p_2) the choice of n (the number of trials made on a level) determines both α and β , or the choice of one of α and β , determines the other and also n .

The discussion of this method has illustrated the considerations which must enter into the choice of a final method. However, this particular one is very inefficient for the task at hand because it requires a large number of trials to complete one determination of the desired percentage point. Now it is a well known empirical result in sampling theory that if one specifies an OC by fixing values of p_1 , p_2 , α and β , then there are many criteria which approximately meet this specification. All criteria which have the same p_1 , p_2 , α and β will, in a sense, estimate a fixed percentage point with the same accuracy, but the average number of trials required will be different for the different criteria. In the present situation, it would be desirable to use the criterion which uses the smallest average number of trials.

b. Sequential Probability Ratio Plan. There is one sampling plan available for use in developing a method which approaches this property of minimum average number of trials. This is the Sequential Probability Ratio Sampling Plan as described by Wald in the Journal of the American Statistical Association, Vol. 40, No. 231, pp. 277-306. The statement is made that this plan has the property that the average number of trials required to reach a decision concerning the location of the desired percentage point is minimized simultaneously at the two levels for which $p_x = p_1$ and p_2 . In general, it will not be possible to obtain a criterion which requires the smallest number of trials for all values of p_x .

The distinguishing characteristic of the Sequential Probability Ratio Sampling Plan is that it does not require a fixed number of trials to reach a decision. Instead it gives a decision on a two way alternative as soon as enough evidence has been accumulated to make the probabilities associated with the two types of error less than or equal to α and β . The application here may be described as follows.

It is desired to estimate the level at which the probability of an explosion is p . On the basis of certain considerations with respect to accuracy and number of trials, values for p_1 , p_2 , α and β will be chosen. p_1 is less than p , and p_2 is greater than p . Once these constants have been chosen, reference to Wald's paper enables one to compute two sequences of integers u_1, u_2, u_3, \dots and d_1, d_2, d_3, \dots . (The subscripts refer to the accumulated number of trials on a particular level.)

Now suppose testing is being done on level x and that it is necessary to decide on the basis of trials whether the level corresponding to p is above x , below x , or is nearly identical with x . As the testing is carried out on this level, a record is made of the trial number (n) and the total number of explosions which have been obtained in these n trials. After each trial the number of explosions is compared with the two sequences above. If, at any point in the testing, the number of explosions in n trials becomes equal to u_n , testing is discontinued and the statement is made that the level corresponding to p is above x . On the other hand, if the number of explosions in n trials becomes equal to d_n , testing is discontinued and the statement is made that the level corresponding to p is below x . As long as neither of these decisions is obtained, testing is continued.

This entire procedure may be stated more precisely as follows:

For each value of n (trial number) we determine

$$u_n = A + B \cdot n \text{ and}$$

$$d_n = C + B \cdot n \text{ where}$$

$$A = \frac{\log \frac{\beta}{1-\alpha}}{\log \frac{p_2}{p_1} - \log \frac{1-p_2}{1-p_1}}$$

$$B = \frac{\log \frac{1-p_1}{1-p_2}}{\log \frac{p_2}{p_1} - \log \frac{1-p_2}{1-p_1}} \quad \text{and}$$

$$C = \frac{\log \frac{1-\beta}{\alpha}}{\log \frac{p_2}{p_1} - \log \frac{1-p_2}{1-p_1}}$$

If u_n is not an integer, we replace it by the largest integer less than u_n . Similarly, if d_n is not an integer, we replace it by the smallest integer which is greater than d_n . Now if we are testing on level x , we continue making an additional trial as long as e_n , the number of explosions in n trials, satisfies $u_n < e_n < d_n$. If for some value n we have $e_n \leq u_n$, then we discontinue testing on level x and move to level $x + \delta$. If for some value n we have $e_n \geq d_n$, then we discontinue testing on level x and move to level $x - \delta$. When the per cent point has been bracketed we stop.

If this procedure is applied to a level x as described above, a decision one way or the other will eventually be reached. The number of trials required to reach this decision will vary from test to test, and may, at times, become quite large. For this reason it has been desirable in the present application to decide upon a maximum number of trials which are to be taken at any one level. If this number of trials is performed on a level with no decision being reached, the statement will be made that the desired percentage point lies on this level. This process of truncation means that the nominal values of α and β are not, in fact, the exact risks. However, for the sake of convenience α and β will be used as referring to the truncated procedure as well as the untruncated procedure in determining the u_n and d_n sequences.

As far as procedure is concerned, there is only one additional step to be considered. If one should start testing on a level quite far removed from the level corresponding to the desired percentage, the Sequential procedure would prove quite costly with respect to total number of trials. For this reason, some one of the simple "staircase" methods should be used to locate a level at which to begin the detailed sequential procedure. Empirical experience with the type of sequential procedure described below suggests that the accuracy of the final level is nearly independent of where the detailed sequential procedure is started.

In order to illustrate the type of design which is obtained from these general considerations, let us suppose that we desire to estimate the 12 per cent point. p_1 will be taken as .08, p_2 as .16, α as .25 and β as .25. Furthermore, no more than thirteen trials will ever be made on a single level. This is the Sequential Plan which is recommended in the general part of this report for the estimation of the 12 per cent point. Then the operator's instructions will read as follows:

- (1). Start at a level where almost no explosions are expected.
- (2). If no explosion is obtained in two trials, move up one step. Continue to move up after each pair of trials until the first explosion occurs.
- (3). After the first explosion, continue to test at this level using the following procedure (disregarding the tests already made):
 - a. If two explosions are obtained out of 2, 3, 4, or 5 trials move down one step as soon as the second explosion is obtained.
 - b. If three explosions are obtained out of 7, 8, 9, 10, 11, 12 or 13 trials move down one step as soon as the third explosion occurs.
 - c. If thirteen trials are made without obtaining an explosion move up one step.
 - d. If no move (as indicated by a, b, or c) has been made at the end of thirteen trials, discontinue testing.
- (4). As long as no decision of type 3(d) is obtained, continue to move up or down as indicated by 3(a), 3(b), or 3(c). Discontinue testing when a decision of type 3(c) follows a decision of type 3(a) or 3(b), or when a decision of type 3(a) or 3(b) follows a decision of type 3(c).
- (5). Record:
 - a. The level at which a decision of type 3(d) has been obtained, or
 - b. The midpoint of the last two levels at which testing occurred when testing has been discontinued as in (4).

The recorded level estimates the approximate 12 per cent point.

It will be noted that this procedure starts with a Single Explosion (two trials on a level) Design.

Investigations have been conducted to see whether certain other interpolation schemes could not be substituted for (5) above which would reduce the variance of estimated percentage point. No scheme was discovered which would make any significant reduction in this variance.

The design pattern for a Sequential Test such as outlined above is determined by fixing values of p_1 , p_2 , α , β , and the maximum number of trials to be used on any one level. In order to obtain some indication of the way in which values should be assigned to these variables, certain computations have been carried out on the

assumption that a cumulative normal distribution represents the probability of explosion as a function of height of test. As in the previous sections, the location of the levels at which tests are to be made will be measured in terms of σ' , the standard deviation of the assumed normal distribution. The point at which 50 per cent explosions are expected will be taken as the origin of this scale of measurement.

In all the computations which follow, p_1 is always taken as .08 and p_2 as .16. In general, increasing p_1 and decreasing p_2 (for fixed α and β) will have the same effect as decreasing α and β .

c. Effect of changes in α and β . The fundamental investigation concerning the choice of α , β , and the point at which truncation occurs will be conducted for an interval size of $.5\sigma'$. Consequently the levels at which testing occurs are $-2.5\sigma'$, $-2.0\sigma'$, $-1.5\sigma'$, $-1.0\sigma'$, $-.5\sigma'$, 0, $+.5\sigma'$, etc. In what follows the σ' will be understood even though it is not explicitly written down. There is no a priori reason for supposing that α should be taken equal to β , but in order to simplify the computations this assumption is made, except for two cases contained in Table 22. Furthermore, each Sequential Scheme will be preceded by a Single Explosion (two trials on a level) Design as given in our illustrative example for $\alpha = \beta = .25$, $p_1 = .08$, $p_2 = .16$, and truncation at 13 trials. For the computations, it was assumed that testing started at $-2.5\sigma'$.

The first set of computations were made for $\alpha = \beta = .15$, $\alpha = \beta = .20$, $\alpha = \beta = .25$ and $\alpha = \beta = .30$. In each instance truncation was made at approximately the same number of trials although some variation was allowed because of the particular characteristics of each plan. In computing the expected value and the variance of the estimated percentage point, the estimate of the percentage point for each repetition of the test was taken as in (5) of the representative plan given in this section. The pertinent data obtained from this investigation are given in Table 22.

TABLE 22

Summary of Results for Sequential Methods
at Interval Size of .5 σ

$$(p_1 = .08, p_2 = .16)$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\alpha = \beta = .15$	33 trials	-1.18	.0738	60.5	9.2	.224	1.47	.889
$\alpha = \beta = .20$	33 trials	-1.16	.0732	56.9	8.4	.240	1.63	.971
$\alpha = \beta = .25$	32 trials	-1.17	.0883	49.7	7.1	.227	1.59	.938
$\alpha = \beta = .30$	34 trials	-1.14	.1134	42.3	6.1	.208	1.45	.854
$\alpha = .20$ $\beta = .30$	30 trials	-1.11	.0907	47.7	7.2	.232	1.53	.921
$\alpha = .30$ $\beta = .20$	33 trials	-1.21	.0884	51.2	7.0	.222	1.61	.935

- (1) Process Truncated at
- (2) Expected Value of Estimated Point
- (3) Variance of Estimated Point
- (4) Average Number of Trials
- (5) Average Number of Explosions
- (6) Accuracy per trial
- (7) Accuracy per Explosion
- (8) Weighted Accuracy

An examination of this table shows that the accuracy per trial, the accuracy per explosion and the weighted accuracy assume maximum values for $\alpha (= \beta)$ between .15 and .30. The variation in these criteria for $\alpha (= \beta)$ between these limits is not very large, and we have chosen to take $\alpha (= \beta) = .25$ for the remainder of the investigation. It is to be noted that as $\alpha (= \beta)$ increases from .15 to .30, the average number of trials and explosions decrease while the variance of the estimated percentage point increases. The expected value of the estimated percentage point varies only from -1.18 (corresponding to 11.9 per cent) to -1.14 (corresponding to 12.7 per cent). If the difference in point of truncation is taken into account, it is seen that the results for $\alpha = .20, \beta = .30$ and $\alpha = .30, \beta = .20$ give approximately the same results as $\alpha (= \beta) = .25$ except that the expected values for the estimated percentage point show slightly more variation.

d. Effect of point of truncation. Let us now examine the effect of truncation on the Sequential Method for $\alpha (= \beta) = .25$. A summary of the results obtained is given in Table 23.

TABLE 23

Effect of Truncation on Sequential Method
at Interval Size of .5%

($\alpha = \beta = .25$, $p_1 = .08$, $p_2 = .16$)

Process Truncated at	Expected Value of Estimated Point	Variance of Estimated Point	Average No. of Trials	Average No. of Explosions	Accuracy per Trial	Accuracy per Explosion	Weighted Accuracy
32 trials	-1.17	.0883	49.7	7.1	.227	1.59	.938
23 trials	-1.16	.0952	43.5	6.4	.240	1.64	.978
13 trials	-1.11	.1249	32.0	5.0	.250	1.60	.976

This table indicates that the accuracy per explosion and the weighted accuracy assume maximum values for truncation between 13 and 32 trials. However, the actual advantage of any one part of this range over any other appears to be very slight. The accuracy per trial is increasing over this range. As one truncates at a smaller number of trials, the average number of trials required for a determination of the percentage point decreases while the variance of the estimate increases. Within this range of truncation the expected value of the estimated percentage point varies from -1.17 (corresponding to 12.1 per cent) to -1.11 (corresponding to 13.4 per cent).

Upon the basis of the data presented in Tables 22 and 23, the general part of this report recommended the use of the Sequential Method with $\alpha (= \beta) = .25$, $p_1 = .08$, $p_2 = .16$ and truncation at 13 trials for the estimation of the 12 per cent point. It is true that truncation at 23 trials would give slightly larger values for the efficiency criteria, but the greater simplicity achieved by truncation at 13 trials outweighs this gain. Operator instructions for this method are given in the general part and are repeated in this section as an illustration.

e. Effect of change in interval size. As a final step we have investigated the stability of this method with respect to changes in interval size. The results obtained are given in Table 24.

TABLE 24

Effect of Interval Size on Sequential Method
 $(\alpha = \beta = .25, p_1 = .08, p_2 = .16)$
 Truncation at 13 Trials

Interval Size	Expected Value of Estimated Point	Variance of Estimated Point	Average No. of Trials	Average No. of Explosions	Accuracy per Trial	Accuracy per Explosion	Weighted Accuracy
.2	-1.19	.1292	40.8	4.1	.190	1.89	.946
.5	-1.11	.1249	32.8	5.0	.250	1.60	.976
1.0	-1.21	.1823	25.6	4.8	.214	1.14	.745

From this table the following observations can be made:

- (1). For interval sizes ranging between $.2\sigma$ and 1.0σ , the expected value of the estimated percentage point varies only from -1.11 (corresponding to 13.4 per cent) to -1.21 (corresponding to 11.3 per cent).
- (2). For maximum accuracy per trial an interval size of about $.5\sigma$ should be used.
- (3). The average number of explosions tends to increase as the step size increases.
- (4). The average number of trials decreases as the step size increases.

For this particular method, it makes very little difference where the testing levels are located with respect to the underlying distribution. Thus for an interval size of 1.0σ , the expected value of the estimated percentage point is -1.21 if testing starts at -2.5σ (test levels being -2.5σ , -1.5σ , $-.5\sigma$, $+.5\sigma$, etc.) and -1.19 if testing starts at -3.0σ (test levels now being -3.0σ , -2.0σ , -1.0σ , 0 , $+1.0\sigma$, etc.).

In this section we have not shown how to compute the results of the use of a Sequential Plan. These details will be discussed in the section on Sequential Methods in the computational portion of this report.

12. The Pursuit Method.

The Pursuit Method was devised to reduce the effect of the assumption of normality. This method can be illustrated by considering its application to the estimate of the 10 per cent point. Its distinguishing characteristics are that

- (i). it requires a predetermined number of trials, and
- (ii). it attempts to concentrate the trials on the two levels between which the 10 per cent point lies.

We proceed as follows:

After a trial has been performed on a level, compute the per cent explosions on this level, taking into account all trials which have been made on this level.

If this percentage is less than 10 per cent, make the next trial on the next higher level.

If this percentage is greater than 10 per cent, make the next trial on the next lower level.

If this percentage is equal to 10 per cent, make the next trial on the same level.

Continue in this fashion until the fixed number of trials have been made.

Estimate the 10 per cent point by using linear interpolation on the two levels which bracket the 10 per cent point.

This method has been investigated by experimental procedures, as explained in Section 24, and the results are summarized in Table 25. Similar experimental procedures can be applied to investigate any proposed method.

TABLE 25

Results^a Obtained by Using the Pursuit Method
to Estimate the 10 per cent Point

(Interval size .50; First trial at -3.0)

Number of Trials	Average No. of Explosions	Average Level	Variance	Accuracy per Trial	Accuracy per Explosion	Weighted Accuracy
25	4.5	-1.14	.1849	.216	1.20	.773
50	7.6	-1.30	.0961	.208	1.37	.826
100	13.6	-1.39	.0576	.174	1.28	.736

* Each value in this table is based upon a set of 40 experimental tests.

For details, see Section 24.

A comparison of the values given in this table for accuracy per trial, accuracy per explosion and weighted accuracy, with those obtained from the other methods under similar conditions, seems to indicate that the Pursuit Method is not as efficient as the others. For example, at interval size .50 the NPFI Method (Table 13) gives .308 for accuracy per trial, 1.60 for accuracy per explosion and 1.06 for weighted accuracy. The Sequential Method (Table 24) gives .250 for accuracy per trial, 1.60 for accuracy per explosion and .976 for weighted accuracy.

13. The Picatinny Method.

The method used at the Picatinny Arsenal differs from the others in one important respect in that the operator is given a large degree of choice in the exact procedure followed. As in all other procedures a set of levels (equally spaced in some scale) is taken. Ten trials are made on each level which is chosen for use in a specific test. The aim of the procedure is to obtain results on two consecutive levels such that there is no explosion on the lower of these two levels and there is at least one on the upper. The upper of these two levels is reported.

An example will illustrate the method. Suppose that the set of levels taken is $-3.0, -2.5, -2.0, \dots, 1.0$. The first ten trials are made at -3.0 and no explosions result. The operator then decides he is below the "10 per cent point" and begins to make ten trials at $+0.5$. When an explosion occurs on the fifth trial, he decides he is above the 10 per cent point. The third set of 10 trials is made at -2.0 and no explosions occur. The operator decides to move up and make ten trials at -1.5 . When an explosion occurs on the eighth trial the test is finished and -1.5 is reported.

It is clear that the number of trials for the test depends on the operator. If he is not skillful in successively bracketing his end point, a very large number of trials will be required. Not only the number of trials, but also the accuracy and the per cent point estimated depend on the exact sequence of operations.

To obtain any numerical results concerning the Picatinny Method it has been necessary to standardize the procedure. It has been assumed that if there is no explosion on the first level at which trials are made then the operator tests on each successive higher level until an explosion occurs. If there is an explosion on the first test level he tests successively on lower levels until he obtains ten non-explosions on a level.

If the test starts at -4.0 and proceeds upward by steps of $.5$, the 7 per cent point is estimated with a variance of $.290$. Since the number of explosions is one, the accuracy per explosion is 3.45 . This value compares favorably with other tests such as NPF (1.61) and the Single Explosion (1.77 at step size 0.5 and 3.40 at 10 trials per σ). The average number of trials is 55.4 , hence, the accuracy per trial is extremely small, namely $.062$. If the test were started at -2.5 about 29.9 trials would be saved, but the variance would change only slightly. Then the accuracy per trial would be $.135$.

An alternative standardization of the procedure is to start where there is a negligible possibility of 10 successive non-explosions and then proceed downwards. Starting at $+0.5$ and using steps of $.5$ one estimates the 13 per cent point, and the variance is $.471$. The number of trials required on the average is 22.1 . The accuracy per trial is $.096$. This figure is lower than when one starts at -2.5 . The average number of explosions is 4.26 and the accuracy per explosion is $.499$.

It should be observed that the Picatinny Method, when standardized to start where there are practically no explosions, is simply a Single Explosion Design with 10 trials per level. As pointed out in Section 9 the accuracy per explosion increases as the number of trials per level increases, while the accuracy per trial decreases.

14. The Up and Down Method.

a. Introduction. The Up and Down Design and statistical analysis were devised for the purpose of estimating the 50 per cent point and the standard deviation (σ) of the assumed underlying normal distribution. Since the normal distribution is completely specified by the 50 per cent point (or mean) and standard deviation, it is obvious that one can estimate any percentage point on the basis of the estimates of the 50 per cent point and the standard deviation. In the present section we shall study the accuracy of this natural method of estimation.

b. The method of estimation. In assuming a normal curve, we assumed

$$P_X = q\left(\frac{X - \mu}{\sigma}\right),$$

so that knowledge of μ and σ would allow the calculation of the x corresponding to given p from

$$x = \mu + k\sigma,$$

where

$$p = q(k).$$

The Up and Down Method^a produces an estimate m of μ and an estimate s of σ , so the natural estimate of x is

$$m + ks.$$

Values of $k + 5.000$ are given in Fisher and Yates, Statistical Tables, under the name of "Probits".

^a AMP Report No. 101.1R indicates how to obtain these estimates.

c. Accuracy. Sampling error in our estimate of the p per cent point arises because of the sampling error in m and s. Since the sampling errors of m and s depend on the step size as well as on the number of trials, our estimate of the p per cent point does too. Figure 18 indicates the standard deviation of m (σ'_m) for a sample of 50; Figure 19 gives the standard deviation of s (σ'_s) for a sample of 50. For example, if the step size is 1.5 (in σ') then σ'_m is .215 and σ'_s is .256.

Since the estimates of μ and σ' are nearly statistically independent the standard deviation of $m + k \cdot s$ is approximately

$$\sigma'^2_m + k^2 \cdot \sigma'^2_s ,$$

where σ'_m and σ'_s are the standard deviations of m and s, respectively. From this expression it is clear that the standard deviation of $m + k \cdot s$ increases with k; that is, the farther from the 50 per cent point the p per cent point is, the greater is the sampling error. For a sample size of 50 and interval size of 1.5 the standard deviation is

$$\sqrt{.0462 + k^2(.0655)} .$$

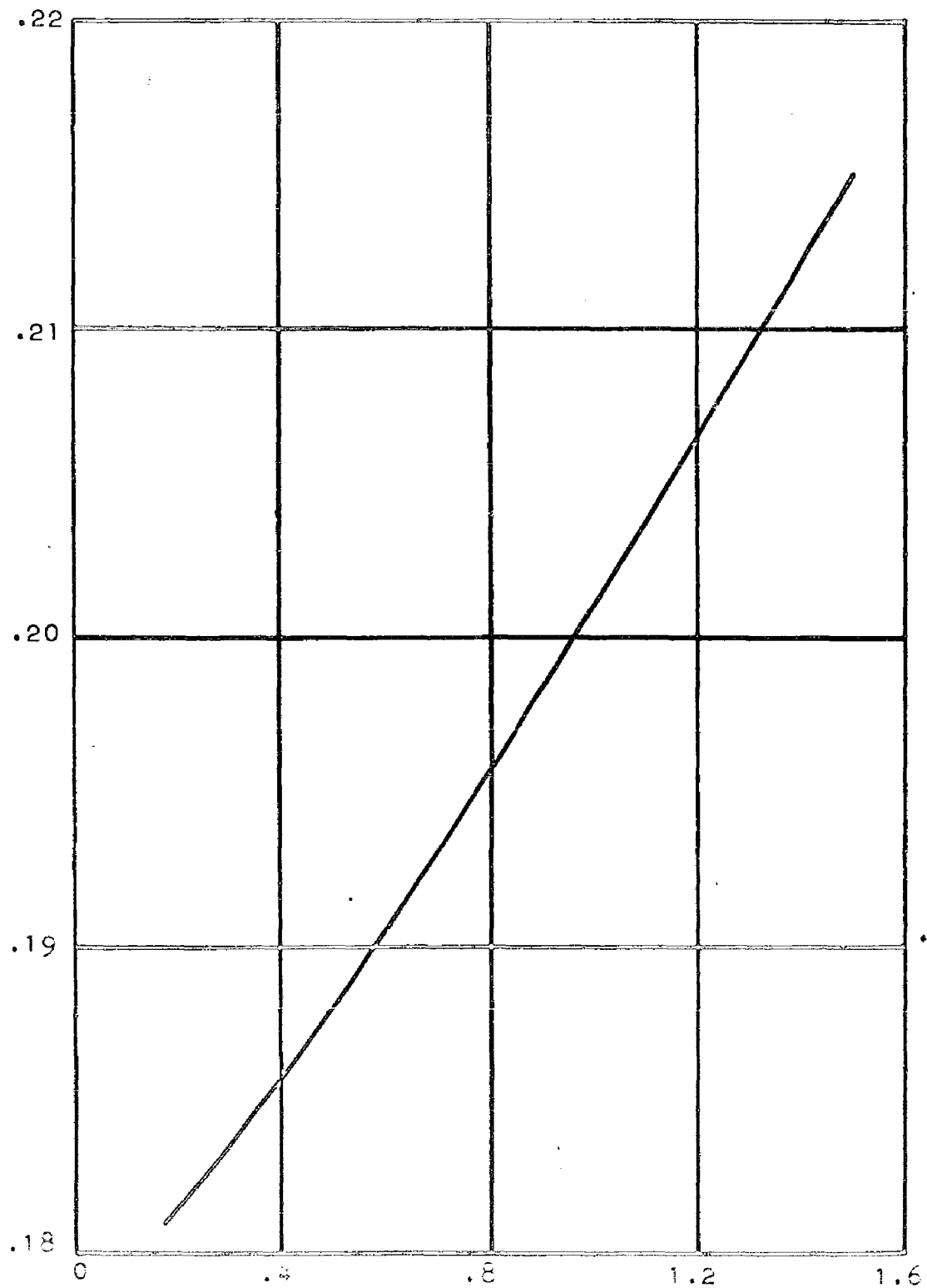
For any sample size (and interval size of 1.5) the accuracy per trial is $1 / (2.312 + k^2 \cdot 3.276)$. This expression is graphed against k in Figure 20. The accuracy per trial for the estimation of the 25 per cent point is .26; the accuracy per trial for the estimation of the 10 per cent point is .13. The latter figure is somewhat lower than the accuracy per trial for some other methods. For per cent points smaller than 10 per cent, this estimate becomes much more inaccurate.

Figure 18

Standard Deviation of the 50 per cent Point
Estimated by the Up and Down Method

Standard Deviation
of the Estimated Mean

(Sample of 50)



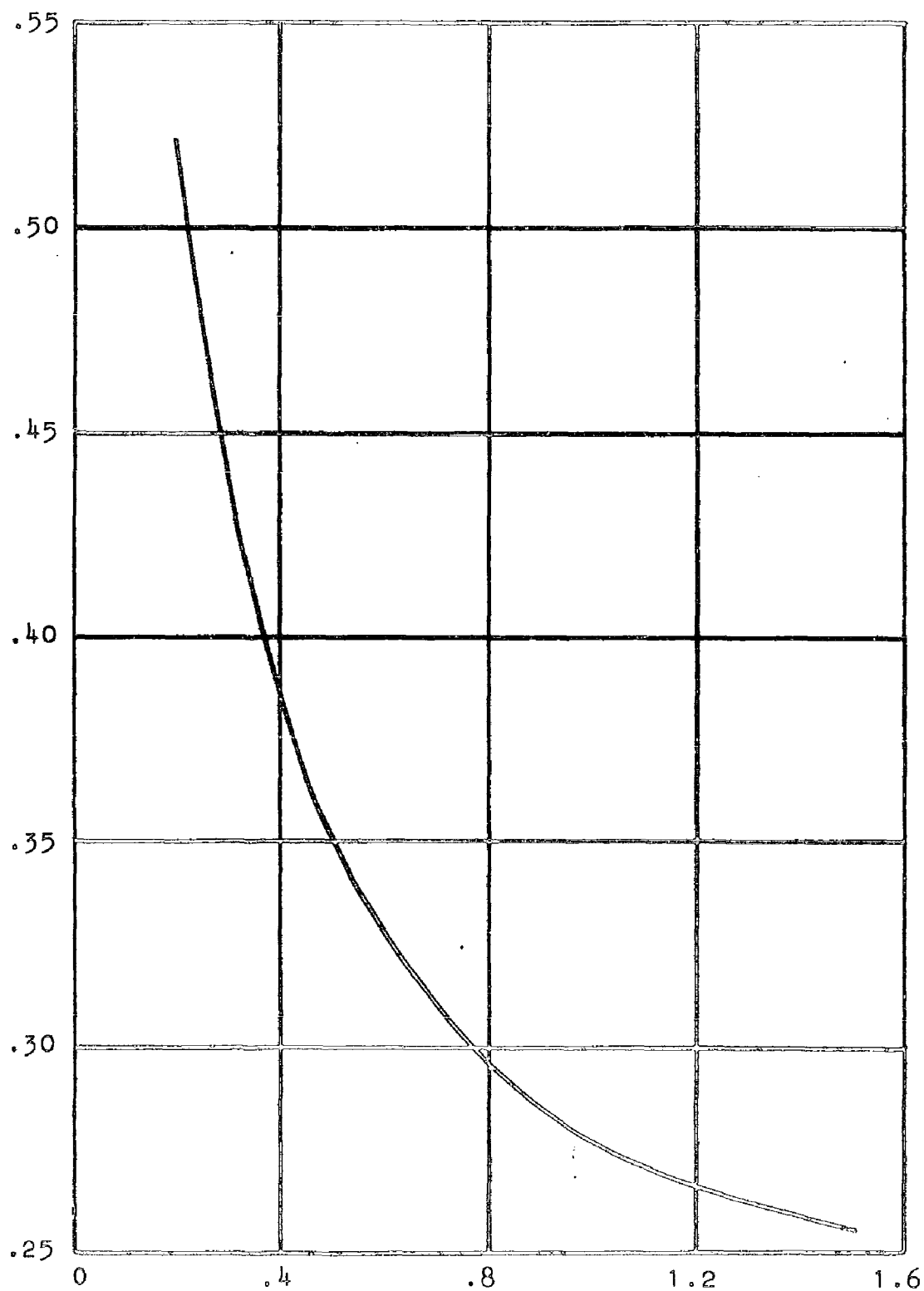
Interval size (δ)

Figure 19

Standard Deviation of the Estimate of the
Population σ by the Up and Down Method

Standard Deviation
of the Estimated σ

(Sample of 50)

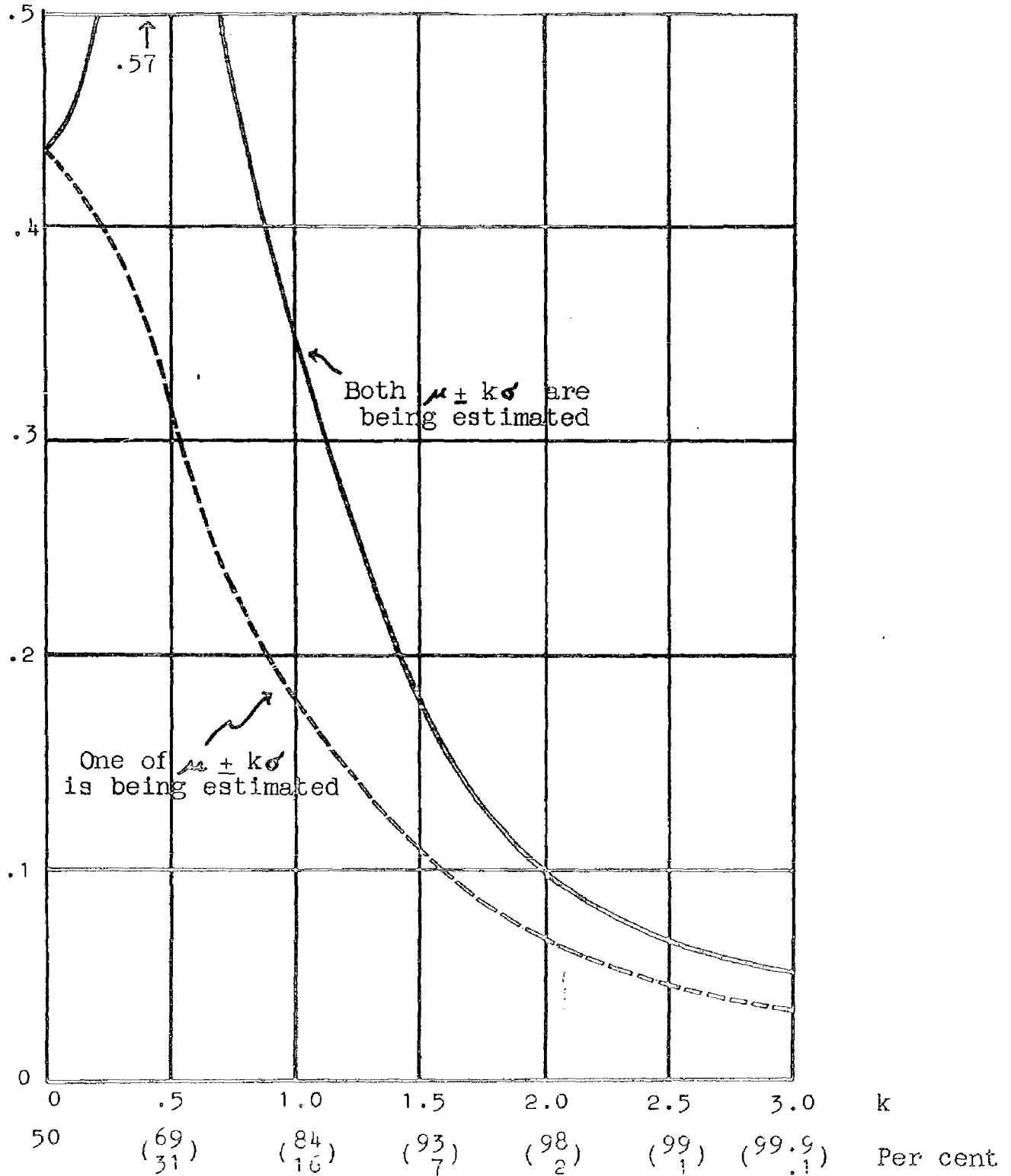


Interval Size (δ)

Figure 20

Accuracy per Trial for Up and Down Method
(Interval size of 1.5σ)

Accuracy per trial



Also included on Figure 20 is a curve which represents our estimate of the accuracy per trial for any sample size and for interval size of 1.5, when one considers that $m \pm k \cdot s$ are estimated simultaneously. The equation of this curve is given by

$$\frac{2 - r^2}{2.312 + k^2 \cdot 3.276},$$

where r is the correlation between $m + k \cdot s$ and $m - k \cdot s$. r is given by the expression

$$\frac{1 - k^2 \frac{\sigma_s'^2}{\sigma_m^2}}{1 + k^2 \frac{\sigma_s'^2}{\sigma_m^2}},$$

On the average one half of the trials will be explosions. Hence, the accuracy per explosion will be twice the accuracy per trial, namely, $1 / (1.156 + k^2 \cdot 1.638)$.

This method is peculiar in that the maximum accuracy per explosion is at the 50 per cent point, where it is .865. However, at the 25 per cent point (accuracy per explosion .519) it is more economical in explosions to use some other method.

Since the weighting of σ_m' and σ_s' depends on the value of k the best step size for estimating $m + k \cdot s$ varies with k . One of these standard deviations (σ_m') increases with increasing step size, while the other (σ_s') decreases. Therefore, the accuracy per trial is increased for small k if the interval size is decreased. Table 26 gives the accuracy per trial for four different step sizes and 4 values of k .

TABLE 26

Accuracy per Trial for Up and Down Method

k \ Step Size	.2	.5	1.0	1.5
0.0	.608	.566	.455	.433
.5	.199	.305	.336	.319
.7	.121	.211	.257	.255
1.0	.066	.128	.172	.179

No larger step size than 1.5 is taken because beyond that point σ_m and σ_s depend on the relation between μ and the levels at which the testing is carried out. If the true mean falls midway between two testing levels, σ_m increases as δ increases, while σ_s decreases slightly. However, if the true mean falls on a level both standard deviations increase. Since one would seldom be certain of the relationship between μ and the testing level, it is better to use a step size no greater than 1.5.

Table 26 shows that beyond a k of about .75 it is more efficient to use a step size of 1.5 than any smaller although the difference between the accuracy per trial for 1.5 and 1.0 is very slight over the range $k = .5$ to $k = 1.0$. In general, for small values of x accuracy per trial is quite insensitive to changes in step size. Roughly speaking, the step size indicated is as follows:

- (1). To estimate per cent points from 44 per cent to 56 per cent use a step of about .2.
- (2). To estimate per cent points from 37 per cent to 44 per cent and 56 per cent to 63 per cent use a step of about .5.
- (3). To estimate per cent points from 24 per cent to 37 per cent and 63 per cent to 76 per cent use a step of about 1.0.
- (4). To estimate per cent points less than 24 per cent and greater than 76 per cent use a step size of about 1.5 and no larger.

For per cent points near 50 per cent one can obtain greater efficiency than indicated by Figure 20. Table 26 shows the improvement in efficiency in terms of accuracy per trial by choosing smaller step sizes when per cent points near 50 per cent are to be

estimated. The maximum increase in accuracy per trial over the values given in Figure 20 is 27 per cent (at the 50 per cent point).

d. Confidence intervals. From Figures 18 and 19 one can compute σ_m and σ_s for a sample of 50 and hence the standard deviation of $m + k \cdot s$ ($=y$, say) which is

$$\sqrt{\sigma_m^2 + k^2 \cdot \sigma_s^2}.$$

If the sample is of size N instead of 50 multiply the standard deviation for 50 by $\sqrt{50/N}$ to obtain the corresponding standard deviation for a sample of size N . For reasonably large samples we may regard y as normally distributed with mean of the per cent point estimated and standard deviation as computed (say σ_y). For smaller samples, however, we can make a correction for lack of normality (see AMP Report No. 101.1R, pp. 20, 21). The statement about confidence intervals can be made as follows: The true value of the number estimated by y will, on the average, lie between the two numbers

$$y - 1.96 \frac{N+2.4}{N} \sigma_y$$

and

$$y + 1.96 \frac{N+2.4}{N} \sigma_y$$

95 times out of 100. The 99 per cent confidence interval is $(y - 2.58 \frac{N+5.2}{N} \sigma_y, y + 2.58 \frac{N+5.2}{N} \sigma_y)$.

e. Comparison with other methods. This method of estimation depends strongly on the assumption of normality. The farther from 50 per cent is the per cent point estimated, the greater is the dependence on this assumption.

The Up and Down Method can as easily be used to estimate both a high and low per cent point as to estimate merely one. Hence, it may be more efficient to use this than one of the other methods when one wishes to estimate two or more points.

The efficiency for one point is rather low. For example, the accuracy per trial for an estimate of $-.7$ (in terms of σ) is about .26. On the other hand the

Single Explosion Method with step size .7 estimates approximately the same point (when properly corrected for interval size) with an accuracy per trial of slightly less than .33.

Suppose that one estimates -1.0 and 1.0 by these two methods (using the Single Explosion Method inverted for 1.0). The accuracy per trial for the Up and Down method is .35 and for the Single Explosion Method (interval size .5) is .16. If the points are farther away from 0, and only one point is desired, then methods other than the Up and Down are better. For example, for the 10 per cent point the accuracy per trial for the Up and Down is about .13 but for the NPF is about .16.

15. Future Research in Methods.

Since it is not possible to foresee all the useful developments of the future, the purpose of this section can only be to outline some apparently promising directions of study. In some cases it is easy to point out exactly what computations need to be done, while in others we can only indicate a goal. The following paragraphs indicate as far as possible goals, methods and motivations.

a. Number of trials. Any staircase method is well adapted to sensitivity testing where (i) preparation of the sample is easy and (ii) the time to conduct (and observe) one trial is short. However, such methods as the Up and Down are equally well applicable to situations where preparation is lengthy and where, for efficiency, a chosen number of samples should be prepared at once. This convenience can be associated either with an advance knowledge of the number of trials which a test will require or with an efficient analysis of tests which have been terminated at a fixed number of trials. Clearly the development of highly efficient tests for the 10 per cent and 90 per cent points where the number of trials, at least within very narrow limits, can be predicted in advance, or where the tests can be efficiently analyzed when terminated at an arbitrary point, would be desirable. Work under f, g, h and i below will be relevant.

b. Corrections for curvature. The methods recommended in this report suffer most from curvature of the per cent point estimated as a function of interval size. The present types of adjustment deal with curvature only at the end of the test, and rather crudely. The following questions require attention:

(i). Can better procedures to correct for curvature be devised?

(ii). Can simple methods be found which are affected less by curvature than the methods of this report?

c. Confidence intervals. When the recommended methods are applied to a case where the assumption of normality holds, it is possible to compute confidence intervals. These are only known for the Up and Down Method. More study here would be useful.

d. Ways of testing normality. How can the assumption of normality be efficiently tested? Comparison of results of the NPF Inverted and the Up and Down Method for the 10 per cent point has been suggested. How good is such a test? How should its results be interpreted? Are there better ways?

e. Non-normal operation. How do the means and variabilities of the recommended methods change when the sensitivity curve is not normal? Some information exists, but not nearly enough.

f. Block Up and Down Methods. Various Up and Down Methods utilizing blocks of trials and not single trials can be easily devised. For example,

(i). make 3 trials at a level, go up if 0 or 1 explosions, go down if 2 or 3 explosions,

(ii). make trials at a level, going up at the first non-explosion, going down at the third explosion (proposed by NPF).

We need to know the means and accuracies per trial associated with several such methods, and also the best methods of analysis.

g. Variable Interval Methods. The possibility of using smaller and smaller intervals as the trials proceed clearly saves a few trials which would otherwise be lost in finding the approximate level desired. This is particularly clear for Cascade Methods. Are there other gains? How much do they amount to?

h. Multiple Cascade Methods. In this report, most Cascade Methods are applied for one or two up sequences. The prominent exception is the Up and Down Method which can be regarded as a Cascade Method with a large (20 or more) variable number of up and down sequences. Intermediate cases need to be studied to assess the effect of starting over on efficiency.

i. Other Sequential Methods. Clearly Sequential Methods for other per cent points and with other accuracies deserve study.

- j. Investigation of boundaries. Section 6 contains some informal hunches as to the permanent limitations of staircase methods. If possible, these hunches should be replaced by knowledge. It is probable that detailed information about many more individual methods is needed before a general study can be profitably completed.
- k. Differential effects of interval size. In this report, effects of interval size have been assessed by calculating means for, say, three interval sizes, drawing a curve and determining a correction graphically. As noted in Section 8 it is easy to compute a differential correction numerically (i. e., one which holds for δ between δ_0 and $\delta_0 + d\delta$), and it seems likely that the use of such tangential estimates would help in understanding the situation and in solving problem j.
- l. Variable length-accuracy. A problem of both practical and theoretical interest is posed by the methods, such as the NPF, which involve a variable number of trials. When few trials are required, is the result more or less accurate than on the average?
- m. Delayed Staircase Methods. In many cases of sensitivity testing, e. g., heat stability, it is necessary to wait a substantial time for the conclusion of the test. Delayed staircase methods would be applicable, e. g., and Up and Down Method where the level for the $N + 1$ th trial depended on the result of the N th trial.
- n. Variable number of trials per level. What would be the characteristics of Cascade Methods in which the number of trials per level increased by a certain number for each higher level at which testing occurred?

III. COMPUTATION

16. Introduction.

The quantities associated with the distribution of a sensitivity test result which are relevant to our analysis are (1) the expected level reported, (2) the variance of the reported level, (3) the average number of trials, and (4) the average number of explosions. In many cases, of course, other quantities were also computed.

17. General Assumptions.

All computations are based on the assumption that the sensitivity curve is a cumulative normal curve. The natural scale of measurement of the levels is in terms of the standard deviation of this normal distribution. We have modified this assumption to the extent that at -3.0 (and lower) the probability of an explosion is taken to be zero and at $+3.0$ (and higher) the probability of an explosion is taken to be unity. For our purposes this modification is unimportant, for relatively few trials in any test procedure would be made at -3 or lower, or at $+3$ or higher, and hence, the distributions of test procedures with the modification of the cumulative normal curve differs little from that without the modification. The computations are simplified by assuming a finite range. We shall let p_x be the probability of an explosion on level x , and q_x be the probability of a non-explosion ($=1-p_x$).

Most of the test procedures require starting the test at a level where the probability of an explosion (or in some cases, of a non-explosion) is nearly zero. In our computations we have generally started at -2.5 (or $+2.5$).

18. Single Explosion Design.

Perhaps the simplest design to treat is the Single Explosion Design with single trials on a level. The probability of the explosion occurring at level x is the product of the probability of a trial being made at x , and the probability that a trial at level x results in an explosion. For example, the probability of the explosion occurring at -2.5 is $.00621$, at -2.0 is $(1 - .00621) \cdot .02275 = .02261$, and at -1.5 is $(1 - .00621) \cdot (1 - .02275) \cdot .06681 = .06488$. Table 27 gives the probability, $P(x)$, of the level at which the explosion occurs with step sizes of $.50$.

TABLE 27

Probability Distribution for the
Single Explosion Design

Level	p_x	q_x	Probability, $P(x)$
-2.5	.00621	.99379	.006
-2.0	.02275	.97725	.023
-1.5	.06681	.93319	.065
-1.0	.15866	.84134	.144
-0.5	.30854	.69146	.235
0.0	.50000	.50000	.264
0.5	.69146	.30854	.182
1.0	.84134	.15866	.068
1.5	.93319	.06681	.012
2.0	.97725	.02275	.001
2.5	.99379	.00621	.000

From this distribution one computes the expected value ($-.2410$) and the variance ($.565$) of the reported level, where $E(x) = \sum_x x P(x)$ and $\sigma_x'^2 = \sum_x x^2 P(x) - [E(x)]^2$.

Because the number of trials depends only on where the explosions occur, the average number of trials, $T(x)$, is a simple function of the expected level, namely

$$T(x) = \frac{E(x) + 2.5}{\delta} + 1.$$

In this case it is 5.518. The average number of explosions is one. From these four numbers one can easily compute the accuracy per trial ($.321$), the accuracy per explosion (1.77), and the weighted accuracy (1.14).

The computation is slightly more complicated in case one makes 2 trials at each level instead of one. Then the probability of at least one explosion at a level \underline{x} is $p_2(x) = 1 - q_x^2$, where q_x is the probability of a non-explosion in a trial at level \underline{x} . This is the same as the probability that one obtains at least one explosion out of two trials where the second trial is made regardless of whether the first is an explosion. Then the probability, $P_2(x)$, of the explosion occurring at \underline{x} is

$p_2(x) \cdot Q_2(x)$, where $Q_2(x)$ is the probability that a trial is made on level x . For example, the probability of an explosion in 2 trials at -2.5 is $1 - (.99379)^2 = .01238$ and the probability of an explosion out of 2 trials at -2.0 is $1 - (.97725)^2 = .04498$. Hence, the probability of the explosion in a test occurring at -2.5 is .01238 and at -2.0 is $.04498 \cdot (1 - .01238) = .04442$. Table 28 gives the probabilities for the explosion levels.

TABLE 28

Probability Distribution for the Single Explosion
Method with 2 Trials on a Level

Level	$p_2(x)$, or $1 - q_x^2$	q_x^2	Probability $P_2(x)^*$
-2.5	.01238	.98762	.012
-2.0	.04498	.95502	.045
-1.5	.12916	.87084	.122
-1.0	.29215	.70785	.240
-0.5	.52188	.47812	.303
0.0	.75000	.25000	.209
0.5	.90480	.09520	.063
1.0	.97483	.02517	.006
1.5	.99554	.00446	.000
2.0	.99948	.00052	.000
2.5	.99996	.00004	.000

* The probabilities were rounded to 3 decimal places and then adjusted to sum to one.

The expected level is $-.657$ and the variance is $.435$. The average number of trials, $T_2(x)$, can be thought of as the sum of the average number of trials on the level of the explosion and the average number of trials on the levels without the explosion. The latter is related to the average level by the formula

$$2 \cdot \frac{E(x) + 2.5}{8}.$$

Evaluating this gives 7.372 . The number of trials on the level of the explosion is one or two depending on whether or not the explosion occurs on the first or second trial. This average is therefore

$$\sum_x P_2(x) \cdot (1 \cdot p_x + 2 \cdot q_x p_x) / (1 - q_x^2) = \sum_x P_2(x) \frac{1 + 2q_x}{1 + q_x}.$$

Evaluating this formula gives 1.406 .

The sum of these two numbers ($7.372 + 1.406 = 8.778$) is the average number of trials. The average number of explosions is one. The accuracy per trial is $.262$, the accuracy per explosion is 2.30 , and the weighted accuracy is 1.22 .

For a Single Explosion Design with k trials per level the computations are made in a similar fashion. Here the probability of at least one explosion out of k trials at a level x is $1 - q_x^k = p_k(x)$. Then the probability of obtaining the explosion at x is $p_k(x) \cdot Q_k(x) = P_k(x)$, where $Q_k(x)$ is the probability that trials are made on level x . The expected number of trials at the level of the explosion is

$$\begin{aligned} \sum_x P_k(x) (1 \cdot p_x + 2 \cdot q_x p_x + 3 \cdot q_x^2 p_x + \dots + k \cdot q_x^{k-1} p_x) / (1 - q_x^k) \\ = \sum_x P_k(x) \frac{1 + 2q_x + \dots + kq_x^{k-1}}{1 + q_x + \dots + q_x^{k-1}}. \end{aligned}$$

* This can also be written as

$$\sum_x P_k(x) \frac{k \cdot q_x^{k+1} - (k+1) q_x^k + 1}{p_x (1 - q_x^k)}.$$

This formula is easy to use if the same computation is done for successive values of k , for then the numerators and denominators are built up step by step.

The expected number of trials at the levels where there is no explosion is simply

$$k \frac{E(x) + 2.5}{\delta}$$

The sum of this and the preceding number is the average number of trials, $T_k(x)$. The average number of explosions is again one. The accuracy per trial, accuracy per explosion, and weighted accuracy follow easily from the above computations.

For this Single Explosion Design there is no possibility of making an adjustment (for fixed δ) to minimize the variance of the reported level since we end the test with only one piece of information, namely the level of the first explosion. However, we can devise an adjustment to make the average reported final level nearly constant with respect to changes in δ within certain limits. The procedure for doing this is discussed in Sections 8 and 10.

19. Cascade Methods.

As an example of the computations involved in the Cascade Methods let us consider the particular method which consists of a Single Explosion Design (one trial at each level) followed by another Single Explosion Design which starts three levels below the level at which the first explosion occurs ($k = 1$, $m = 1$, $h = 3$). The first run starts at -2.5 and ends at level x . The second run starts at level $x - 1.5$ ($\delta = .5$) and ends at level y . The probability of an explosion at y starting at a level, say -1.0 , is written $_{-1.0}P(y)$ (then $P(x)$ as used previously would be written $_{-2.5}P(x)$). Thus

$$_{-1.0}P(y) = q_{-1.0} \cdot q_{-.5} \cdots q_{y-.5} \cdot p_y$$

Table 29 gives the probability of ending at level y for starting points from -2.5 to $+3.0$. Note that, since the probability of an explosion below level -2.5 is 0, if one starts the second run lower than -2.5 , the probability that the run results in an explosion at level y is $_{-2.5}P(y)$.

TABLE 29

Probability Distribution for a Single Explosion
Run Starting at Various Levels

(one trial on each level)

Each Entry Represents ${}_wP(y)$

$y \backslash w$	-2.5	-2.0	-1.5	-1.0	-.5	0	.5	1.0	1.5	2.0	2.5	3.0
-2.5	.0062											
-2.0	.0226	.0226										
-1.5	.0649	.0653	.0668									
-1.0	.1438	.1447	.1481	.1587								
-0.5	.2353	.2367	.2422	.2596	.3085							
0.0	.2636	.2653	.2714	.2909	.3457	.5000						
0.5	.1823	.1834	.1877	.2011	.2391	.3457	.6915					
1.0	.0684	.0689	.0705	.0755	.0897	.1298	.2596	.8413				
1.5	.0120	.0121	.0124	.0133	.0158	.0228	.0457	.1481	.9332			
2.0	.0009	.0009	.0009	.0009	.0011	.0016	.0032	.0104	.0653	.9772		
2.5	0	0	0	0	0	0	.0001	.0002	.0015	.0226	.9938	
3.0	0	0	0	0	0	0	0	0	0	.0001	.0062	1.0000

* The factors p_x and q_x used in computing this table will be found in Table 27.

As described in the Technical Part, Section 8, the reported level is to be a function of the difference of the two end levels (i. e., of $y-x = r$). It proves to be expedient to first compute a table in which the columns show the probabilities, $P(x,y)$, for a constant value of y , of the first run ending at levels of x from -2.5 to +3.0 and the second run ending at the particular level, y . For example, the first column is obtained by computing

$$P(-2.5,y) = {}_{-2.5}P(-2.5) {}_{-4.0}P(y) = {}_{-2.5}P(-2.5) {}_{-2.5}P(y)$$

for $y = -2.5, -2.0, \dots, +3.0$, and the fifth column by computing

$$P(-.5, y) = {}_{-2.5}P(-.5) {}_{-2.0}P(y)$$

for the same \underline{y} values as above. Summing the rows gives the probability, $P(\underline{y})$, of ending at level \underline{y} . Table 30 gives the values of $P(x, \underline{y})$ for the Cascade Method $k = 1, m = 1, h = 3$.

TABLE 30

Probability Distribution for the Cascade Design,
 $k = 1, m = 1, h = 3$

Each Entry Represents the Probability
of the First Run Ending at Level x,
and the Second Run Ending
at Level y, i. e., $P(x,y)$

[illegible]

Another table is then constructed by rearranging the first table so that the columns show the probabilities $P(x,r)$ for a given r and varying values of x . Summing the columns gives $P(r)$. The values of $P(x,r)$ and $P(r)$ are given in Table 31. Reference to Section 8 on the possible adjustments shows that we can modify the reported level, for each value of r , so that the average reported level is the same regardless of the value of r . Under these circumstances the variance of the reported level will be equal to the sum of the x variances for each value of r , each variance being weighted by the appropriate probability. That is,

$$\text{Variance} = \sum_r P(r) \left[\sum_x \frac{x^2 P(x,r)}{P(r)} - \left(\sum_x x \frac{P(x,r)}{P(r)} \right)^2 \right]$$

For $k = 1$, $m = 1$, and $h = 3$, the variance is .2921.

The number of trials required to reach a final level, y , does not depend on the end point of the first run. This follows from the fact that the number of trials in the first run is

$$1 + \frac{x - (-2.5)}{\delta} + 1$$

and in the second run is

$$1 + \frac{y - (x - h\delta)}{\delta} + 1.$$

If we add these two together we eliminate x , and obtain for the total number of trials (i. e., for fixed y)

$$\frac{y + 2.5}{\delta} + h + 2.$$

From this expression, we obtain the average number of trials to be

$$\sum_y P(y) \frac{y + 2.5}{\delta} + h + 2.$$

In the case $k = 1$, $m = 1$, $h = 3$, the average number of trials is 9.6650. The number of explosions is two. It follows that the accuracy per trial is .354, the accuracy per explosion is 1.71, and the weighted accuracy is 1.15.

TABLE³¹

Probability Distribution for the Cascade Design,
 $k = 1, m = 1, h = 3$

Each Entry Represents the Probability
 of the First Run Ending at Level x ,
 and the Second Run Ending
 at Level $x+r$, i. e., $P(x,r)$

End Level of the First Run, x	Value of r											
	-1.5	-1.0	-.5	0	.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
-2.5					.0001	.0004	.0009	.0015	.0016	.0011	.0004	.0001
-2.0			.0001	.0005	.0015	.0032	.0053	.0060	.0041	.0015	.0003	
-1.5		.0004	.0015	.0042	.0093	.0153	.0171	.0118	.0044	.0008	.0001	
-1.0	.0009	.0032	.0093	.0207	.0338	.0379	.0262	.0098	.0017	.0001		
-0.5	.0054	.0154	.0341	.0557	.0624	.0432	.0162	.0028	.0002			
0.0	.0176	.0390	.0638	.0715	.0495	.0186	.0033	.0002				
0.5	.0289	.0473	.0530	.0367	.0138	.0024	.0002					
1.0	.0211	.0236	.0164	.0061	.0011	.0001						
1.5	.0060	.0041	.0016	.0003								
2.0	.0006	.0002										
2.5	0	0										
3.0	0											
$P(r)$.0805	.1332	.1798	.1957	.1715	.1211	.0692	.0321	.0120	.0035	.0008	.0001

* This Table is obtained by a rearrangement of Table 30.

The corrections to be made to the result of each test in order to minimize the variance of this reported level and to make its expected value independent of changes in δ are discussed in Sections 8 and 10.

Now let us consider the situation where k trials are made on each level for the first run and m trials are made on each level for the second run (h given). Then, as described in the $k = 1, m = 1$ case, we can compute the probability $P_{k,m}(x,y)$ of the first run ending on level x and the second run ending on level y . If we sum these probabilities (for a fixed x) over all values of y we obtain $P_k(x)$, and if we sum them (for fixed y) over all values of x we obtain $P_{k,m}(y)$. Finally, we can sum them for all values of y and x for which $(y-x)$ is a constant r and so obtain $P_{k,m}(y-x) = P_{k,m}(r)$. Using $P_{k,m}(x,r)$ we can compute the minimum variance of the average reported level, the computations proceeding as in the $k = 1, m = 1$ case.

The average number of trials required to obtain particular values of x and y is equal to the sum of the average number of trials required to reach level x plus the average number of trials required to go from level x to y . These two expressions are, from the section on the Single Explosion Design and the earlier results of this section, equal to

$$k \frac{x+2.5}{\delta} + \frac{1 + 2q_x + \dots + kq_x^{k-1}}{1 + q_x + \dots + q_x^{k-1}}, \text{ and}$$

$$m \frac{y-(x-h\delta)}{\delta} + \frac{1 + 2q_y + \dots + mq_y^{m-1}}{1 + q_y + \dots + q_y^{m-1}}.$$

Consequently the average number of trials is simply

$$\sum_{x,y} P_{k,m}(x,y) \left[k \frac{x+2.5}{\delta} + \frac{1 + 2q_x + \dots + kq_x^{k-1}}{1 + q_x + \dots + q_x^{k-1}} \right. \\ \left. + m \frac{y-(x-h\delta)}{\delta} + \frac{1 + 2q_y + \dots + mq_y^{m-1}}{1 + q_y + \dots + q_y^{m-1}} \right].$$

The best method for computing this quantity will depend upon the values of k and m , and the previous computations which have been made. For example, we find that when $k = m = 1$, $k(\frac{x+2.5}{\delta}) + m(\frac{y-(x-h\delta)}{\delta}) = (\frac{y+2.5}{\delta} + h)$. Moreover the sum of the other two terms is 2 so that the above expression for the average number of trials reduces to

$$\sum_y P_{1,1}(y) \cdot \frac{y+2.5}{\delta} + h + 2$$

It will usually be advantageous, since we have to introduce $(y-x)$ in order to compute the variance, to write $m(\frac{y-(x-h\delta)}{\delta})$ as $m(\frac{y}{\delta} + h)$ and use the probabilities $P_{k,m}(r)$ to compute this portion of the expression.

The number of explosions is always 2 so we can now readily compute accuracy per trial, accuracy per explosion, and weighted accuracy.

20. Single Explosion plus m Trials.

This is an extension of the Single Explosion Design (Section 18). The simplest case ($k = 1$, $m = 1$) is to make single trials on successively higher levels until an explosion occurs at level x and then to make one more trial at level $x-h\delta$. Starting at -2.5 the probability of an explosion at level x and an explosion at level $x-h\delta$ is expressed by

$$P(x) p_{x-h\delta}$$

where $P(x)$ is the same as $_{-2.5}P(x)$ of the last section. Similarly the probability of an explosion at level x and of a non-explosion at level $x-h\delta$ is given by

$$P(x) q_{x-h\delta}$$

As an example consider the probability of obtaining an explosion on the run up at level -1.5 and another explosion at level -1.0 ($h = -1$, $\delta = .5$). $P(-1.5)$ is .06488 and $p_{-1.0}$ is .15866, so the probability is $.06488 \cdot .15866 = .010$. The probability of obtaining an explosion at level -1.5 and a non-explosion at level -1.0

is $.06488 \cdot .84134 = .055$. In this manner a distribution is computed showing the probability of the single trials resulting in an explosion at level \underline{x} and the one additional trial resulting in an explosion at level $x-h\delta$. Similarly a distribution is computed for the probability of a non-explosion at level $x-h\delta$. Table 32 shows these values for $h = -1$ and $\delta = .5$.

TABLE 32

Probability Distributions for a Single Explosion
plus \underline{m} Trials Design
($k = 1, m = 1, h = -1$)

Level \underline{x}	Probability of an explosion at \underline{x} , and at $x + .5$	Probability of an explosion at \underline{x} , and of a non-explosion at $x + .5$
-2.5	0	.006
-2.0	.002	.021
-1.5	.010	.055
-1.0	.044	.100
-0.5	.117	.118
0.0	.182	.082
0.5	.153	.029
1.0	.063	.005
1.5	.012	0
2.0	.001	0
2.5	0	0

The expected level of the x 's in the cases resulting in explosions at level $x-h\delta$ is determined by evaluating

$$\frac{\sum_x x P(x) p_{x-h\delta}}{\sum_x P(x) p_{x-h\delta}}$$

For our example it is .065.

The expected level of the x 's in the cases resulting in non-explosions at level $x-h\delta$ is determined by evaluating

$$\frac{\sum_x x P(x) q_{x-h\delta}}{\sum_x P(x) q_{x-h\delta}}.$$

For our example it is $-.671$.

As in the preceding section on Cascade Methods, we are going to adjust the final level so that the average reported level of those tests which end in an explosion is the same as the average reported level for the tests which end in a non-explosion. Under these circumstances the variance of the reported level is

$$\begin{aligned} & \sum_x x^2 P(x) p_{x-h\delta} - \frac{\left(\sum_x x P(x) p_{x-h\delta}\right)^2}{\sum_x P(x) p_{x-h\delta}} \\ & + \sum_x x^2 P(x) q_{x-h\delta} - \frac{\left(\sum_x x P(x) q_{x-h\delta}\right)^2}{\sum_x P(x) q_{x-h\delta}}. \end{aligned}$$

The variance in the particular example we have been discussing is $.434$.

The expected number of trials required will be one more than the expected number of trials required in the Single Explosion Design (one trial on a level), that is

$$\frac{E(x) + 2.5}{\delta} + 2.$$

For $\delta = .5$ and $h = -1$, this value is 6.518 . Similarly the average number of explosions is

$$1 + \sum_x P(x) p_{x-h\delta}$$

or 1.584. Consequently the accuracy per trial is .353, the accuracy per explosion is 1.45, and the weighted accuracy is 1.03.

Now let us consider the case where we make k trials on a level at successively higher levels until an explosion results, and then take trials on level $x-h\delta$ either until an explosion occurs or m non-explosions. The probability is $P_k(x)$ that we obtain our first explosion on level x on the up sequence, $q_{x-h\delta}^m$ that we obtain m non-explosions on level $x-h\delta$ and $1-q_{x-h\delta}^m$ that we obtain an explosion on level $x-h\delta$. Consequently the probability is

$$P_k(x) q_{x-h\delta}^m$$

that we obtain our first explosion on level x and no explosions on level $x-h\delta$, and

$$P_k(x) \cdot [1 - q_{x-h\delta}^m]$$

that we obtain our first explosion on level x and an explosion on level $x-h\delta$.

The variance in this case is computed as in the case where $k = 1$, $m = 1$ with the substitution of $P_k(x) q_{x-h\delta}^m$ for $P(x) q_{x-h\delta}$ and $P_k(x) [1 - q_{x-h\delta}^m]$ for $P(x) p_{x-h\delta}$.

The average number of trials required is equal to

$$T_k(x) + \sum_x P_k(x) \left[\sum_{i=1}^m 1 - q_{x-h\delta}^{i-1} p_{x-h\delta} + m q_{x-h\delta}^m \right],^*$$

where $T_k(x)$ is defined in Section 18.

* This expression is readily computed if we are dealing with successively increasing values of m .

This expression can also be written as

$$T_k(x) = \sum_x P_k(x) \frac{(1 - q_{x-h\delta}^m)}{p_{x-h\delta}}.$$

21. NPF Inverted Method.

In the NPF Inverted Method one makes a single trial at successively higher levels until an explosion occurs at level \underline{x} , and then at level $x - \delta$ one starts making three trials on a level at successively lower levels until all three trials are non-explosions. (One stops testing at a level as soon as an explosion occurs and starts again at the next lower level.) The run up starts at a level where the probability of an explosion is almost zero. For $\delta = .5$ this would be at -2.5 .

The probability of the first run ending at \underline{x} is $P(x)$. The probability of the second run ending at \underline{y} , if one starts at $x - \delta$, is the probability that there is an explosion at each of the levels $x - \delta$, $x - 2\delta$, $x - 3\delta$, ... $y + \delta$, and that there are three non-explosions at \underline{y} . This probability is $(1 - q_{x-\delta}^3) \cdot (1 - q_{x-2\delta}^3) \dots (1 - q_{y+\delta}^3) \cdot q_y^3$. Then this product times the probability of an explosion at \underline{x} is denoted by $Q(x, y)$. For example, if $x = 1.0$, $P(x) = .144$ from Table 27 and $Q(-1.0, -2.0) = .144 \cdot (1 - (.93319)^3) \cdot (.97725)^3 = .02514$. A table is then computed in which the columns show the probability, $Q(x, y)$, of the run down ending at varying levels of \underline{y} and the run up ending at a given \underline{x} . The rows show the probabilities, $Q(x, y)$, of the run up ending at varying levels of \underline{x} and the run down ending at a given \underline{y} . Table 33 gives the values of $Q(x, y)$. Summing the rows gives the probability of ending at \underline{y} , $Q(y)$, regardless of where the run up ends. The table is rearranged, so the columns are the probabilities $Q(y, r)$, for a constant difference in end levels ($x - y = r$), and for varying values of \underline{y} . Summing the columns gives the probabilities, $Q(r)$, of \underline{r} , regardless of the final end level. $Q(y, r)$ and $Q(r)$ are given in Table 34.

The expected end level of the second run $E(y|r)$, is computed for each value of \underline{r} by evaluating

$$E(y|r) = \frac{\sum y Q(y, r)}{Q(r)}$$

TABLE 33

Probability Distribution for the NPFI Method

Each Entry^{*} Represents the Probability
of the Run up Ending at Level \underline{x} ,
and the Run Down Ending
at Level \underline{y} , $Q(x,y)$

End Level of Run Down, \underline{y}	$Q(\underline{y})$	End Level of the Run up, \underline{x}									
		-2.5	-2.0	-1.5	-1.0	-0.5	0	.5	1.0	1.5	2.0
-3.0	.00679	.00621	.00042	.00008	.00003	.00002	.00002	.00001			
-2.5	.03100		.02219	.00425	.00176	.00117	.00088	.00053	.00019	.00003	
-2.0	.12562			.06055	.02514	.01664	.01248	.00755	.00275	.00048	.00003
-1.5	.30246				.11685	.07733	.05800	.03509	.01279	.00224	.00016
-1.0	.33630					.14011	.10509	.06358	.02317	.00406	.00028
-0.5	.16270						.08715	.05273	.01921	.00337	.00023
0.0	.03264							.02279	.00830	.00145	.00010
0.5	.00238								.00201	.00035	.00002
1.0	.00005									.00005	

* For $x = 2.5$ the entries for $y = -1.0$ and -0.5 are both .00001 and are zero for all other values of y .

TABLE 34

Probability Distribution for the NPFI Method

Each Entry Represents the Probability
of the Run Down Ending at Level y ,
and the Run Up Ending at
Level $y+r$, $Q(y,r)$

End Level of Run Down, y	Value of r							
	.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
-3.0	.00621	.00042	.00008	.00003	.00002	.00002	.00001	
-2.5	.02219	.00425	.00176	.00117	.00088	.00053	.00019	.00003
-2.0	.06055	.02514	.01664	.01248	.00755	.00275	.00048	.00003
-1.5	.11685	.07733	.05800	.03509	.01279	.00224	.00016	
-1.0	.14011	.10509	.06358	.02317	.00406	.00028	.00001	
-0.5	.08715	.05273	.01921	.00337	.00023	.00001		
0.0	.02279	.00830	.00145	.00010				
0.5	.00201	.00035	.00002					
1.0	.00005							
$Q(r)$.45791	.27361	.16074	.07541	.02553	.00583	.00085	.00006

Summing these expected end levels over r gives for the average expected end level

$$E(y) = \sum_r Q(r) \cdot E(y|r) .$$

In accordance with our procedure for correcting y so that its variance is a minimum, the variance of the reported level is

$$\sum_r Q(r) \left[\sum_y y^2 \frac{Q(y,r)}{Q(r)} - \left(\sum_y y \frac{Q(y,r)}{Q(r)} \right)^2 \right]$$

The average number of trials is the sum of the average number of trials made to reach \underline{x} and the average number to reach \underline{y} . The average number made to reach \underline{x} is

$$\sum_x \left(\frac{x+2.5}{6} + 1 \right) \cdot P(x) .$$

The number of trials on each level of the run down is one, two or three depending upon whether the first or second trial is an explosion. Consequently the expected number of trials on a level \underline{y} , given that we actually make trials on this level, is equal to

$$1 \cdot p_y + 2 \cdot q_y p_y + 3 q_y^2 (q_y + p_y) .$$

Now the probability that we test on level \underline{y} , and that the run up ends on level \underline{x} , is

$$\frac{Q(x, y)}{q_y^3}$$

From these two expressions we obtain the average number of trials required on the run down as

$$\begin{aligned} \sum_{x, y} \frac{Q(x, y)}{q_y^3} \left[1 \cdot p_y + 2 q_y p_y + 3 q_y^2 (q_y + p_y) \right] \\ = \sum_y Q(y) \frac{1 - q_y^3}{1 - q_y} \cdot \frac{1}{q_y^3} \end{aligned}$$

Thus the average number of trials required for the NPF Inverted is

$$\sum_x \left(\frac{x+2.5}{\delta} + 1 \right) \cdot P(x) + \sum_y Q(y) \frac{1-q_y^3}{1-q_y} \cdot \frac{1}{q_y^3} .$$

The correction for variance and interval size for this method are made in a fashion similar to that described in Sections 8 and 10.

The number of explosions for the run up is one, and for the run down it is one for each level above the final one. This is

$$\frac{x-\delta-y}{\delta} = \frac{r}{\delta} - 1 .$$

The average number of explosions is then

$$1 + \sum_r Q(r) \left(\frac{r}{\delta} - 1 \right) = \sum_r \frac{r}{\delta} \cdot Q(r) .$$

Having determined the variance, trials and explosions, the three accuracy functions are easily found.

22. The Picatinny Method.

The computations for the Picatinny Method are similar to those for the Single Explosion Designs. The Picatinny computations have been done with five decimal place accuracy. With one standardization the test starts at -4.0. Hence the usual truncation of the cumulative normal distribution was not made. For example, the probability of 10 successive non-explosions at -4.0 is .99970. The same probabilities were used for the procedure starting at +.5. The two distributions are given in Table 35.

TABLE 35

Probability Distribution for the Picatinny Method

	Start at -4.0	Start at +.5
-4.0	.00030	.00000
-3.5	.00230	.00007
-3.0	.01338	.00490
-2.5	.05943	.07722
-2.0	.19007	.31762
-1.5	.36664	.40115
-1.0	.30251	.17310
-0.5	.06374	.02495
0.0	.00163	.00098
+0.5	.00000	.00001

To simplify the computations for the procedure starting at -4.0, the set of 10 non-explosions at the lower level was not required if the explosion occurred at -4.0. Similarly in the case of starting at +.5 the explosion was not required if the set of 10 non-explosions occurred at +.5. The error caused by these simplifications cannot be more than .03 per cent and .001 per cent, respectively.

In this case the average number of trials was computed by summing the expected number on each level (regardless of whether an explosion results). The expected number on level x (when trials are made on that level) is

$$\frac{1 - q_x^{10}}{1 - q_x} = \frac{p_{10}(x)}{p_x}$$

If the test starts at -4.0, the expected number on level x is

$$q_{10}(-4.0) q_{10}(-4.0 + \delta) \dots q_{10}(x - \delta) \frac{p_{10}(x)}{p_x},$$

where $q_{10}(y) = 1 - p_{10}(y)$ is the probability that no explosion occurs in 10 trials at level y . If the test starts at $+5$, the expected number of trials on level x is

$$\frac{p_{10}(+.5) p_{10}(+.5-\delta) \dots p_{10}(x)}{p_x}$$

The number of explosions is one when the tests starts at -4.0 . In the other case it is

$$\frac{.5 - E(x)}{\delta} + 1$$

(5.261 for $\delta = .5$).

23. Sequential Method.

In this section we shall show how to obtain the characteristics of a sequential plan set up as described in Section 11. In computing the OC, average number of trials, etc., it has been found inadvisable to use the approximation given by Wald. This arises primarily from the fact that we are truncating, but it is also true that some of his approximations, especially for average sample size, do not seem to be sufficiently accurate in the range in which we choose to take p_1 , p_2 , α and β .

To make the situation clearer, let us consider a specific case, namely $p_1 = .08$, $p_2 = .16$, $\alpha (= \beta) = .25$, truncation at 13 trials. Using the formulas of Section 13, the sequences given in Table 36 for u_n and d_n are obtained.

TABLE 36

Critical Number of Explosions as a
Function of Trial Number

n	u_n	d_n
	*	
1	-	-
2	-	2
3	-	2
4	-	2
5	-	2
6	-	-(3)**
7	-	3
8	-	3
9	-	3
10	-	3
11	-	3
12	-	3
13	0	3

* - means that no move can be made at these positions.

** It is impossible to move down on trial 6 since one would need to have had 2 explosions at some preceding trial, and the move down would have been made then.

Now let us assume that we are testing at level x , the probability of an explosion being p_x and that of a non-explosion, $q_x (=1-p_x)$. The quantities which are necessary in order to carry out an analysis are:

- (1). $p(x, \delta)$, the probability of moving to level $x+\delta$,
- (2). $p(x, -\delta)$, the probability of moving to level $x-\delta$,
- (3). $p(x, 0)$, the probability of no move from level x ,
- (4). The average number of trials and the average number of explosions, given that a particular one of the above results is obtained, namely $T(x, \delta)$, $T(x, -\delta)$, $T(x, 0)$, $D(x, \delta)$, $D(x, -\delta)$ and $D(x, 0)$.

It is obvious from Table 36 that $p(x, \delta) = q_x^{13}$, $T(x, \delta) = 13$ and $D(x, \delta) = 0$. However, the remainder of the expressions do not follow so readily. For example, the probability that we make a decision to move from x to $x-\delta$ on the 9th trial is equal to $K(9,3) q_x^6 p_x^3$, where $K(9,3)$ is the number of ways of obtaining 6 non-explosions and 3 explosions in 9 trials so that no decision to move down has been made before nine trials. In order to obtain the values of the K factors, sometimes called path-coefficients, it is convenient to consider Table 37.

TABLE 37

Path Coefficients for $\alpha (= \beta) = .25$, $p_1 = .08$, $p_2 = .16$

Trial No.	Number of Explosions				
	0	1	2	3	4
1	1	1			
2	1	2	<u>1</u>		
3	1	3	<u>2</u>		
4	1	4	<u>3</u>		
5	1	5	<u>4</u>		
6	1	6	5		
7	1	7	11	<u>5</u>	
8	1	8	18	<u>11</u>	
9	1	9	26	<u>18</u>	
10	1	10	35	<u>26</u>	
11	1	11	45	35	
12	1	12	56	<u>45</u>	
13	<u>1</u>	(13)	(68)	<u>56</u>	
14		13	81	68	
15		13	94	149	<u>68</u>
16		13	107	243	<u>149</u>
17		13	120	350	<u>243</u>
18		13	133	470	<u>350</u>
19		13	146	603	<u>470</u>
20		13	159	749	<u>603</u>

This table is built up by successive addition, and each entry represents the number of ways of obtaining a definite number of explosions (0, 1, 2, 3, or 4) out of a fixed number of trials. If a test on a particular level results in a combination of trials and explosions whose position in the table is underlined, then a decision is made to either move up or down. Consequently, the numbers underlined are the required path coefficients, and these numbers are not used in building up that portion of the table lying below their particular position. For example, $K(8,2) = K(7,1) + K(7,2)$ while $K(8,3) = K(7,2)$ since on all tests having 3 explosions out of 7 trials, we must move down one level, and no more trials are necessary. For purposes of illustration this table has been extended beyond 13 trials by making use of the extension of Table 36.

From this table of path-coefficients one may readily compute

$$p(x, \delta) = q_x^{13}$$

$$p(x, 0) = 13 q_x^{12} p_x + 68 q_x^{11} p_x^2$$

$$\begin{aligned} p(x, -\delta) = & p_x^2 + 2q_x p_x^2 + 3q_x^2 p_x^2 + 4q_x^3 p_x^2 \\ & + 5q_x^4 p_x^3 + 11q_x^5 p_x^3 + 18q_x^6 p_x^3 + 26q_x^7 p_x^3 \\ & + 35q_x^8 p_x^3 + 45q_x^9 p_x^3 + 56q_x^{10} p_x^3 . \end{aligned}$$

Values of $p(x, \delta)$, $p(x, 0)$, and $p(x, -\delta)$ are given in Table 38.

TABLE 38

Values for $p(x, \delta)$, $p(x, 0)$, and $p(x, -\delta)$

(Interval Size = .5)

x	$p(x, \delta)$	$p(x, 0)$	$p(x, -\delta)$
-3.0	1.000000	0	0
-2.5	.922210	.077364	.000426
-2.0	.741437	.251708	.006876
-1.5	.407033	.520664	.072307
-1.0	.105842	.515402	.378818
-0.5	.008260	.159750	.831983
0.0	.000122	.009887	.989995

The required mean values for number of trials and number of explosions can be readily obtained by use of the separate terms making up $p(x, \delta)$, $p(x, 0)$ and $p(x, -\delta)$. Thus

$$p(x, 0) \cdot D(x, 0) = 1 \cdot (13q_x^{12} p_x) + 2 \cdot (68q_x^{11} p_x^2)$$

$$D(x, \delta) = 0$$

$$\begin{aligned} p(x, -\delta) \cdot D(x, -\delta) = & 2 \cdot (p_x^2 + 2q_x p_x^2 + 3q_x^2 p_x^2 + 4q_x^3 p_x^2) \\ & + 3 \cdot (5q_x^4 p_x^3 + 11q_x^5 p_x^3 + 18q_x^6 p_x^3 + 26q_x^7 p_x^3) \\ & + 35q_x^8 p_x^3 + 45q_x^9 p_x^3 + 56q_x^{10} p_x^3) \end{aligned}$$

$$T(x, 0) = 13$$

$$T(x, \delta) = 13$$

$$\begin{aligned}
p(x, -\delta) T(x, -\delta) &= 2(p_x^2) + 3(2q_x p_x^2) + 4(3q_x^2 p_x^2) + 5(4q_x^3 p_x^2) \\
&+ 7(5q_x^4 p_x^3) + 8(11q_x^5 p_x^3) + 9(18q_x^6 p_x^3) + 10(26q_x^7 p_x^3) \\
&+ 11(35q_x^8 p_x^3) + 12(45q_x^9 p_x^3) + 13(56q_x^{10} p_x^3).
\end{aligned}$$

Let us now assume that we start the sequential testing at level \underline{s} . Then the probability that our recorded level is \underline{x} (a level on which trials are made) is $P(x|s)$, where

$$P(x|s) = p(s, -\delta) \cdot p(s-\delta, -\delta) \dots p(x+\delta, -\delta) \cdot p(x, 0)$$

where \underline{x} is less than \underline{s} . If \underline{x} is equal to \underline{s} , we have

$$P(x|s) = p(s, 0),$$

and if \underline{x} is greater than \underline{s} .

$$P(x|s) = p(s, \delta) \cdot p(s+\delta, \delta) \dots p(x-\delta, \delta) \cdot p(x, 0).$$

Similarly, the probability that the recorded level is $x+.5\delta$, given that sequential testing is started at level \underline{s} , is

$$P(x+.5\delta | s) = p(s, -\delta) p(s-\delta, -\delta) \dots p(x+\delta, -\delta) p(x, \delta) \quad x < s$$

$$P(x+.5\delta | s) = p(s, \delta) p(s+\delta, -\delta) \quad x = s$$

$$P(x+.5\delta | s) = p(s, \delta) p(s+\delta, \delta) \dots p(x, \delta) p(x+\delta, -\delta) \quad x > s$$

Table 39 gives the values of these various probabilities for the particular example which we are considering. The values given for -3.0 , -2.5 , -2.0 , etc. refer to $P(x|s)$ while those given for -2.75 , -2.25 , -1.75 , etc. refer to $P(x+.5\delta | s)$, since δ is equal to $.5$.

Now if we precede the sequential testing by a Single Explosion (two trials on a level) Design starting at -2.5δ , the probability that we start sequential testing on level \underline{s} is $P_2(s)$. The computation of $P_2(s)$ has been explained in Section 18.

TABLE 39

Probability Distribution for the Sequential Method

 $(\alpha = \beta = .25, p_1 = .08, p_2 = .16)$

Truncation at 13 Trials, Interval Size of .5

Each Entry Represents the Probability that a
Particular Level will be Recorded if
Sequential Testing Starts at Level s

Recorded Level	Level at which Sequential Method is Started, s						
	-2.5	-2.0	-1.5	-1.0	-0.5	0.0	0.5
-3.00	0	0	0	0	0	0	0
-2.75	.000426	.000003	0	0	0	0	0
-2.50	.077364	.000532	.000038	.000015	.000012	.000012	.000012
-2.25	.006341	.006341	.000458	.000173	.000145	.000143	.000143
-2.00	.232128	.251708	.018200	.006895	.005736	.005679	.005679
-1.75	.049441	.053611	.053611	.020309	.016897	.016728	.016728
-1.50	.356010	.386040	.520664	.197237	.164098	.162456	.162456
-1.25	.105430	.114323	.154191	.154191	.128285	.127001	.127001
-1.00	.143443	.155543	.209786	.515402	.428806	.424515	.424515
-0.75	.024508	.026575	.035843	.088059	.088059	.087178	.087178
-0.50	.004706	.005103	.006882	.016908	.159750	.158152	.158152
-0.25	.000241	.000261	.000352	.000865	.008177	.008177	.008177
-0.00	.000002	.000003	.000004	.000009	.000082	.009887	.009887

Consequently the probability that the final recorded level is x is given by

$$\sum_s P_2(s) P(x|s),$$

where s takes on all values for which $P_2(s) \neq 0$. Similarly the probability that the final recorded level is $x+.5$ is

$$\sum_s P_2(s) \cdot P(x+.5|s).$$

The values of these probabilities are given in Table 40.

TABLE 40

Probability Distribution for the Sequential
Testing Combined with a Single Explosion
(2 trials on a level) Design

Recorded Level	Probability
-3.00	.000000
-2.75	.000005
-2.50	.000967
-2.25	.000536
-2.00	.021059
-1.75	.024154
-1.50	.227164
-1.25	.136417
-1.00	.406228
-0.75	.077976
-0.50	.097708
-0.25	.005024
-0.00	.002776

Finally, the expected value of the estimated percentage point becomes

$$\sum_x \left[x \sum_s P_2(s) \cdot P(x|s) + (x+.5\delta) \sum_s P_2(s) \cdot P(x+.5\delta | s) \right].$$

The variance of the estimated percentage point can be calculated by the use of the same probabilities and is equal to

$$\sum_x \left[x^2 \sum_s P_2(s) P(x|s) + (x+.5\delta)^2 \sum_s P_2(s) P(x+.5\delta | s) \right] - \left\{ \sum_x \left[x \sum_s P_2(s) P(x|s) + (x+.5\delta) \sum_s P_2(s) P(x+.5\delta | s) \right] \right\}^2$$

The treatment of average number of trials and average number of explosions can be explained in a similar fashion. For example, if we start sequential testing at level \underline{s} and record level \underline{x} (where \underline{x} is less than \underline{s}), the total average number of trials required for this will be

$$T(x|s) = T(s, -\delta) + T(s-\delta, -\delta) + \dots + T(x+\delta, -\delta) + T(x, 0).$$

In general, we obtain the total average number of trials corresponding to $P(x|s)$ or $P(x+.5\delta|s)$, by substituting $T(\quad)$ for $p(\quad)$ and replacing the multiplications by additions. $T(x, 0)$, $T(x, \delta)$ and $T(x, -\delta)$ are given in Table 41, while $T(x|s)$ and $T(x+.5\delta|s)$ for the values of \underline{x} and \underline{s} pertinent to our example are given in Table 42. Table 43 gives $D(x, 0)$, $D(x, -\delta)$ and $D(x, \delta)$; Table 44 gives $D(x|s)$ and $D(x+.5\delta|s)$.

Denote by $T_2(s)$ the average number of trials required when we obtain the first explosion on level \underline{s} by use of a Single Explosion (two trials on a level) Design starting at -2.5 . Then the total average number of trials required when sequential testing is started on level \underline{s} is

$$T_2(s) + T(x|s).$$

Consequently the total average number of trials required for the sequential method is

$$\sum_s \sum_x (T_2(s) + T(x|s)) \cdot P_2(s) \cdot P(x|s) \\ + \sum_s \sum_x (T_2(s) + T(x+.5\delta|s)) \cdot P_2(s) \cdot P(s+.5\delta|s) \dots$$

A similar expression is obtained for the average number of explosions by replacing $T(\quad)$ by $D(\quad)$. Note that in this instance $D_2(s) = 1$.

As explained in Section 11, no correction is needed in the Sequential Method to minimize the variance of the average reported level or to adjust for changes in the interval size $\underline{\delta}$.

TABLE 41

Values for $T(x, \delta)$, $T(x, 0)$ and $T(x, -\delta)$
(Interval Size .50)

x	$T(x, \delta)$	$T(x, 0)$	$T(x, -\delta)$
-3.0	13.00		4.75
-2.5	13.00	13.00	5.99
-2.0	13.00	13.00	7.20
-1.5	13.00	13.00	7.41
-1.0	13.00	13.00	6.25
-0.5	13.00	13.00	4.26
0.0	13.00		2.60
0.5			

TABLE 42

Average Number of Trials Required when Sequential
Testing Starts at Level s and the Recorded
Level is x or $x+.5\delta$

Recorded Level x or $x+.5\delta$	Level at which Sequential Testing is Started, s					
	-2.5	-2.0	-1.5	-1.0	-0.5	0.0
-3.00						51.46
-2.75					44.60	46.71
-2.50				38.35	39.85	46.71
-2.25			30.94	33.60	39.85	40.72
-2.00		23.74	26.19	27.61	33.86	40.72
-1.75		18.99	20.20	27.61	33.86	33.52
-1.50		18.99	20.20	20.41	26.66	33.52
-1.25		26.00	13.00	20.41	26.66	26.11
-1.00		33.20	26.00	13.00	19.25	26.11
-0.75		39.00	33.41	26.00	19.25	19.86
-0.50		46.41	39.00	32.25	13.00	19.86
-0.25		52.00	45.25	39.00	17.26	15.60
0.00		58.25	52.00	43.26	26.00	
		65.00	56.26	52.00		
		69.26	65.00			
		78.00				

24. The Pursuit Method.

Although this method has not been recommended for general use, it seems desirable to indicate the procedures which were used in determining its characteristics. The primary reason for doing this is explained by the fact that it is extremely difficult to calculate these characteristics exactly as was done for the other methods, and one is forced to adopt an experimental procedure. This experimental procedure has wide applicability and often allows one to obtain results which would be impossible to realize in any other manner.

The essential feature of this procedure is that we conduct a large number of simulated sensitivity experiments, and then calculate the mean and variance of the estimated percentage point from the resulting tests. This mean and variance are of course subject to sampling error and so we do not obtain the exact value of the population parameters. However, we can decrease this sampling error to any desired value by conducting enough simulated tests. For our work on the Pursuit Method, a sample of 40 tests was used in each instance.

In conducting such a simulated experiment we desire to determine, when testing on a particular level, whether each trial is an explosion or non-explosion. Clearly we could accomplish this by having, for this level, a box containing black and white balls in the proper proportions and then drawing the balls from the box in a random fashion. A white ball would represent an explosion and a black ball would represent a non-explosion. As in the other portions of this report, the correct proportions of balls for each level would be determined from the cumulative normal curve (see Figure 1). Another method uses some results of such experiments that have been tabulated in Sankhya (the Indian Journal of Statistics), Vol. I, pp. 303-328, and we can refer directly to these tables. These tables are of random numbers (drawn from a normal distribution).

In order to illustrate the use of these tables, we shall conduct one test, consisting of 20 trials, for the determination of the 10 per cent point. Let us suppose that testing is being done on levels $-3.0, -2.5, -2.0, \dots, 0, .5, \dots$ and that testing starts at -3.0 . A block of 20 values selected from the table in Sankhya

is given below:

+0.157
 -0.615
 +0.736
 +1.163
 +0.371
 -0.316
 -0.336
 +0.556
 +0.399
 +0.568
 +0.787
 -0.705
 +0.367
 +0.120
 +1.723
 +2.616
 +0.725
 +0.021
 -0.435
 +1.105

Our only rule of procedure in using these values is that if testing is being done on level x , then a value from the table less than x represents an explosion in this trial and one greater than x represents a non-explosion. This is equivalent to the assumption that the relation between per cent explosions and level of severity can be represented by a cumulative normal curve. Now following the rules given in Section 14, we obtain the following results for our sample of 20:

Results of an Experimental Test
 with the Pursuit Method

-3.0	o								
-2.5		o							
-2.0			o						
-1.5				o					
-1.0					o		o	o	o
-0.5					o	o	o	x	o
0.0						x	o	o	

o represents a non-explosion

x represents an explosion

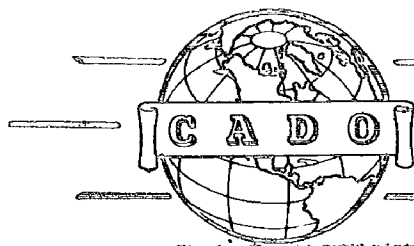
We see that the per cent explosions on level -0.5 is 12 and on level -1.0 is 0. Consequently, using linear interpolation, the level corresponding to the 10 per cent point is $-0.5 + (.16) \cdot (-0.5) = -.58$.

We can now repeat this as many times as desired, obtaining an estimate of the 10 per cent point each time, and then calculate the mean and the variance of this estimate from this sample of values.

DISTRIBUTION LIST

Office of Research and Inventions (5)
 BuShips (1)
 BuAer (1)
 NOL, Navy Yard, Washington, D.C. (6)
 David Taylor Model Basin (2)
 Naval Mine Depot, Yorktown, Virginia (1)
 Commanding Officer, Naval Powder Factory, Indian Head, Md. (1)
 Commanding Officer, Naval Proving Ground, Dahlgren, Va. (1)
 USNA Postgraduate School, Annapolis, Md. (1)
 Naval Ordnance Test Station, Inyokern, Calif. (3)
 ChOrd, War Dept., Washington, D.C., Attn: Safety and Security Branch,
 Technical Unit (2)
 ChOrd, War Dept., Washington, D.C., Attn: SPOTM and Pyro. Section (6)
 ChOrd, War Dept., Washington, D.C., Attn: Research and Development Service,
 Ammunition Development Division, Technical Section (1)
 ChOrd, Field Director of Ammunition Plants, St. Louis 8, Mo. (1)
 Picatinny Arsenal, Dover, New Jersey (2)
 Officer in Charge, Ballistic Research Lab., Aberdeen Proving Ground, Md. (3)
 Officer in Charge, Watertown Arsenal, Watertown, Conn. (1)
 British Admiralty Delegation, 1025 Vermont Ave., Washington, D.C. (3)
 British Commonwealth Scientific Office, Dupont Circle Bldg., Dupont Circle,
 Washington, D.C. (6)
 U. S. Naval Attache', Technical Section, London, England (2)
 Armament Research Dept., Woolwich, England (3)
 OSRD Liaison Office, 1530 P St., N.W., Washington, D.C. (3)
 Eastern Laboratory, Explosives Department, duPont deNemours and Co.,
 Wilmington, Delaware (1)
 Officer in Charge, National Research Council, Ottawa, Canada. Attn:
 Director of Explosives Research Laboratory (1)
 Experiment Station, Hercules Powder Company, Wilmington 99, Delaware. (1)
 Frankford Arsenal, Philadelphia 37, Pennsylvania (1)
 Statistical Research Group, Princeton University, Princeton, New Jersey (2)
 Naval Ammunition Depot, Mare Island, California (1)
 Naval Ammunition Depot, St. Juliens Creek, Virginia (1)
 Naval Ammunition Depot, Crane, Indiana (1)

Reproduced by



CENTRAL AIR DOCUMENTS OFFICE

WRIGHT-PATTERSON AIR FORCE BASE - DAYTON, OHIO

REEL-C

5228

A.T.I

116518

"NOTICE: When Government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the U.S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto."

UNCLASSIFIED